

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 19:12:02 ; Search time 21.6 Seconds
(without alignments)

520.702 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 1095

Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIREICRSSSSSESTG 216

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1084.5	99.0	491	1 MDM2_HUMAN	Q00987 homo sapien
2	1035.5	94.6	491	1 MDM2_HORSE	P56951 equus caball
3	1004.5	91.7	487	1 MDM2_CANPA	P56950 canis famil
4	842.5	76.9	466	1 MDM2_MESAU	Q60524 mesocricetu
5	805	73.5	489	1 MDM2_MOUSE	P23804 mus musculus
6	497	45.4	473	1 MDM2_XENLA	P56273 xenopus lae
7	408.5	37.3	445	1 MDM2_BRARE	O42354 brachydanio
8	277	25.3	489	1 MDM4_MOUSE	O35618 mus musculus
9	252.5	23.1	490	1 MDM4_HUMAN	O15151 homo sapien
10	96.5	8.8	635	1 HEMA_CHICK	P07997 gallus gall
11	92	8.4	1102	1 MYSC_CHICK	P29616 gallus gall
12	90	8.2	1465	1 YH85_SCHPO	O9P5N0 schizosacch
13	89	8.1	1401	1 CND1_HUMAN	Q15021 homo sapien
14	88	8.0	900	1 MUT5_TREPA	O83348 treponema p
15	87.5	8.0	397	1 SDG1_HUMAN	O60524 homo sapien
16	87.5	8.0	664	1 LAMA_HUMAN	P02545 homo sapien
17	87.5	8.0	933	1 RGA4_SCHPO	O74360 schizosacch
18	87.5	8.0	1271	1 RBMG_HUMAN	O9upn6 homo sapien
19	87	7.9	402	1 DCX_HUMAN	O43602 homo sapien
20	87	7.9	500	1 GAR2_SCHPO	P41891 schizosacch
21	87	7.9	888	1 PHFE_HUMAN	O94880 homo sapien
22	87	7.9	1781	1 AK12_HUMAN	O02952 homo sapien
23	87	7.9	2278	1 FAB1_YEAST	P34756 saccharomyc
24	86.5	7.9	361	1 FKB4_SCHPO	O74191 schizosacch
25	86.5	7.9	1417	1 BLM_HUMAN	P54132 homo sapien
26	86.5	7.9	3788	1 LYST_MOUSE	P97412 mus musculus
27	86	7.9	2326	1 CCAB_DISOM	P56698 discopysge o
28	85.5	7.8	337	1 Z265_HUMAN	O95218 homo sapien
29	85.5	7.8	698	1 ALT_EPT6	Q38433 bacterioph
30	85	7.8	386	1 CGB2_ORYIU	Q3d936 oryzias luz
31	85	7.8	1816	1 AF6_HUMAN	P55196 homo sapien
32	85	7.8	1818	1 HMW2_MYCPN	P75471 mycoplasma
33	85	7.8	1938	1 MYHD_HUMAN	O9ukx3 homo sapien

RESULT 1

MDM2_HUMAN
ID MDM2_HUMAN STANDARD; PRT; 491 AA.
AC Q00987; Q13226; Q13297; Q13298; Q13300; Q13301; Q9UGI3;
AC Q9UMT8;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) [p53-binding protein
DE Mdm2] (Oncoprotein Mdm2) (Double minute 2 protein) (Hdm2).
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM MDM2).
RX MEDLINE=92310576; PubMed=1614537;
RA Oliner J.D., Kinzler K.W., Meltzer P.S., George D.L.,
RA Vogelstein B.;
RA "Amplification of a gene encoding a p53-associated protein in human
RT sarcomas.";
RL Nature 358:80-83 (1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS MDM2-A; -B; -C; -D AND -E).
RP TISSUE=Ovarian carcinomas;
RX MEDLINE=96313107; PubMed=8705862;
RA Sigalas I., Calvert A.H., Anderson J.J., Neal D.E., Lunec J.;
RT "Alternatively spliced mdm2 transcripts with loss of p53 binding
RT domain sequences: transforming ability and frequent detection in human
RT cancer.";
RL Nat. Med. 2:912-917 (1996).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM MDM2-ALPHA).
RX MEDLINE=20065171; PubMed=10597303;
RA Veldhoen N., Metcalfe S., Milner J.;
RT "A novel exon within the mdm2 gene modulates translation initiation in
RT vitro and disrupts the p53-binding domain of mdm2 protein.";
RL Oncogene 18:7026-7033 (1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM MDM2).
RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM MDM2).
RC TISSUE=Muscle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

ALIGNMENTS

34	84.5	7.7	377	1	BEXC_HAEIN	P22930 haemophilus
35	84.5	7.7	574	1	LAMC_MOUSE	P11516 mus musculus
36	84.5	7.7	633	1	MLH_TETH	P40631 tetrahymena
37	84.5	7.7	665	1	LAMA_MOUSE	P48678 mus musculus
38	84.5	7.7	665	1	LAMA_RAT	P48679 rattus norv
39	84	7.7	388	1	CGB2_ORYCU	Q9dga3 oryzias cur
40	84	7.7	682	1	ALT_BPT4	P12726 bacterioph
41	84	7.7	1324	1	MSH6_ARATH	O04716 arabidopsis
42	84	7.7	2567	1	M18B_HUMAN	Q8iug5 homo sapien
43	83.5	7.6	240	1	SVP_A3ATH	Q9fvc1 arabidopsis
44	83.5	7.6	740	1	PEC1_PIG	Q95242 sus scrofa
45	83.5	7.6	1704	1	VITI_FUNHE	Q90508 fundulus he

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A., RA Fahney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., RA Blakesley R.W., Trichwood J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallus D.E., RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A., RA "Generation and initial analysis of more than 15,000 full-length RT human and mouse cDNA sequences"; RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002). [16]

RA "SEQUENCE OF 6-491 FROM N.A. (ISOFORM MDM2-A1). RA Liang H., Atkins H., Abdel-Fattah R., Suaeun R., Lunec J.; RA "Genomic organisation of the human MDM2 oncogene and relationship to its alternatively spliced mRNA's"; RA Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases. [17]

RA "SEQUENCE OF 1-24 FROM N.A. RA MEDLINE=95380270; PubMed=7651818; RA Zauberman A., Flusberg D., Haupt Y., Barak Y., Oren M.; RA "A functional p53-responsive intronic promoter is contained within the human mdm2 gene."; RA Nucleic Acids Res. 23:2584-2592 (1995). [18]

RA "SEQUENCE OF 1-9 FROM N.A. RA MEDLINE=97413643; PubMed=9270029; RA Landers J.B., Cassel S.L., George D.L.; RA "Translational enhancement of mdm2 oncogene expression in human tumor cells containing a stabilized wild-type p53 protein."; RA Cancer Res. 57:3562-3568 (1997). [19]

RA "SEQUENCE OF 301-481 FROM N.A. RA MEDLINE=20542019; PubMed=11087894; RA Taubert H., Kappler M., Meve A., Bartel P., Schlott T., RA Lautenschlaeger C., Bache M., Schmidt H., Wuerli P.; RA "A MboI polymorphism in exon 11 of the human MDM2 gene occurring in normal blood donors and in soft tissue sarcoma patients: an indication for an increased cancer susceptibility?"; RA Mutat. Res. 456:39-44 (2000). [10]

RA "MUTAGENESIS OF CYS-464. RA MEDLINE=98111004; PubMed=9450543; RA Honda R., Tanaka H., Yasuda H.; RA "Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53."; RA FEBS Lett. 420:25-27 (1997). [11]

RA "MUTAGENESIS OF CYS-449. RA MEDLINE=20190101; PubMed=10723139; RA Honda R., Yasuda H.; RA "Activity of MDM2, a ubiquitin ligase, toward p53 or itself is dependent on the RING finger domain of the ligase."; RA Oncogene 19:1473-1476 (2000). [12]

RA "MUTAGENESIS. RA MEDLINE=20187618; PubMed=1072742; RA Fang S., Jensen J.P., Ludwig R.L., Vousden K.H., Weissman A.M.; RA "Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself and p53."; RA J. Biol. Chem. 275:8945-8951 (2000). [13]

RA "MUTAGENESIS OF CYS-441 AND CYS-478. RA MEDLINE=20076498; PubMed=10608892; RA Sharp D.A., Kratowicz S.A., Sank M.J., George D.L.; RA "Stabilization of the MDM2 oncoprotein by interaction with the structurally related MDMX protein."; RA J. Biol. Chem. 274:38189-38196 (1999). [14]

RA "NUCLEOLAR LOCALIZATION SIGNAL. RA MEDLINE=20173879; PubMed=10707090; RA Lohrum M.A.E., Ashcroft M., Kubbutat M.H.G., Vousden K.H.;

RT "identification of a cryptic nucleolar-localization signal in MDM2."; RA Nat. Cell Biol. 2:179-181 (2000). [15]

RA "PHOSPHORYLATION BY ATM. RA MEDLINE=20079591; PubMed=10611322; RA Khosravi R., Maya R., Gottlieb T., Oren M., Shiloh Y., Shkedy D.; RA "Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation in response to DNA damage."; RA Proc. Natl. Acad. Sci. U.S.A. 96:14973-14977 (1999). [16]

RA "X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 25-109 IN COMPLEX WITH P53. RA MEDLINE=97081050; PubMed=8875929; RA Kussie P.H., Gorina S., Marechal V., Elenbaas B., Moreau J., RA Levine A.J., Pavletich N.P.; RA "Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain."; RA Science 274:948-953 (1996). [17]

RA "-1- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN. FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2, TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS. CC "-1- COFACTOR: Zinc is required for ubiquitin ligase E3 activity. CC "-1- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOLASTOMA PROTEIN (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION FACTOR. CC "-1- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC. EXPRESSED PREDOMINANTLY IN THE NUCLEOLAR. INTERACTS WITH ARF(P14) RESULTS IN THE LOCALIZATION OF BOTH PROTEINS TO THE NUCLEOLUS. THE NUCLEOLAR LOCALIZATION SIGNALS IN BOTH ARF(P14) AND MDM2 MAY BE NECESSARY TO ALLOW EFFICIENT NUCLEOLAR LOCALIZATION OF BOTH PROTEINS. CC "-1- ALTERNATIVE PRODUCTS: CC Event=Alternative splicing; Named isoforms=8; CC Name=Mdm2; CC IsoId=Q00987-1; Sequence=Displayed; CC Name=Mdm2-A; CC IsoId=Q00987-2; Sequence=VSP_003208; CC Name=Mdm2-A1; CC IsoId=Q00987-3; Sequence=VSP_003208, VSP_003214; CC Name=Mdm2-B; CC IsoId=Q00987-4; Sequence=VSP_003209; CC Name=Mdm2-C; CC IsoId=Q00987-5; Sequence=VSP_003211; CC Name=Mdm2-D; CC IsoId=Q00987-6; Sequence=VSP_003210; CC Name=Mdm2-E; CC IsoId=Q00987-7; Sequence=VSP_003212, VSP_003213; CC Name=Mdm2-alpha; CC IsoId=Q00987-8; Sequence=VSP_003207; CC "-1- TISSUE SPECIFICITY: UBIQUITOUS. ISOFORMS MDM2-A, -B, -C, -D AND -E ARE OBSERVED IN A RANGE OF HUMAN CANCERS BUT ABSENT IN NORMAL TISSUES. CC "-1- INDUCTION: BY DNA damage. CC "-1- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1. REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND ITSELF. CC "-1- PTM: PHOSPHORYLATED IN RESPONSE TO IONIZING RADIATION IN AN ATM-DEPENDENT MANNER. CC "-1- DISEASE: Seems to be amplified in certain tumors (including soft tissue sarcomas, osteosarcomas and gliomas). A higher frequency of splice variants lacking p53 binding domain sequences was found in late-stage and high-grade ovarian and bladder carcinomas. Four of the splice variants show loss of p53 binding. CC "-1- MISCELLANEOUS: MDM2 RING FINGER MUTATIONS THAT FAILED TO UBIQUITINATE P53 IN VITRO DID NOT TARGET P53 FOR DEGRADATION WHEN

Query Match 99.0%; Score 1084.5; DB 1; Length 491;
 Best Local Similarity 99.5%; Pred. No. 1.4e-76;
 Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGA VTTTQIPASQETLVPRKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
 DB 1 MCNTNMSVPTDGA VTTTQIPASQETLVPRKPLLLKLLKSVGAQKDTYTMKEVLFYLQY 60

QY 60 IMTKRLYDEKQOHVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDSGT 119
 DB 61 IMTKRLYDEKQOHVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDSGT 120

QY 120 SVSNRCHLEGGSDQKDLVQELQEPSSSHLVSRPSTSSRRRAISSEENSDLSGERQ 179
 DB 121 SVSNRCHLEGGSDQKDLVQELQEPSSSHLVSRPSTSSRRRAISSEENSDLSGERQ 180

QY 180 RKRHKSISLSPDESALCVIREICERSSSSESTG 216
 DB 181 RKRHKSISLSPDESALCVIREICERSSSSESTG 217

RESULT 2
 MD2_HORSE STANDARD; PRT; 491 AA.

AC P56951;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
 DE Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (Edm2).
 GN MDM2.

OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OC NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20218866; PubMed=10754200;
 RA Nasir L., Burr P.D., McFarlane S.T., Gault E., Thompson H.,
 RA Argyle D.J.;
 RT "Cloning, sequence analysis and expression of the cDNAs encoding the
 RT canine and equine homologues of the mouse double minute 2 (mdm2)
 RT proto-oncogene.";
 RL Cancer Lett. 152:9-13(2000).

CC -1- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
 CC APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
 CC FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
 CC TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
 CC TARGETS IT FOR PROTEASOME-MEDIATED DEGRADATION (BY SIMILARITY).
 CC -1- COFACTOR: Zinc is required for ubiquitin ligase E3 activity (By
 CC similarity).
 CC -1- SUBUNIT: BINDS P53, P73, ARF (P14), RIBOSOMAL PROTEIN L5 AND
 CC SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOBLASTOMA PROTEIN
 CC (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION
 CC FACTOR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. Expressed
 CC predominantly in the nucleoplasm (By similarity).
 CC -1- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
 CC G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
 CC REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
 CC INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
 CC FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
 CC ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
 CC PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
 CC ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
 CC ITSELF (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the MDM2 / MDM4 family.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
 CC -1- SIMILARITY: Contains 1 SWIB domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; AF121140; AAP28866.1; -;
 CC HSP; O9UMT8; 1YCR.
 CC InterPro; IPR003121; SWIB.
 CC InterPro; IPR001876; Znf_RangDP.
 CC InterPro; IPR001841; Znf_Ring.
 CC Pfam; PF02201; SWIB; 1.
 CC Pfam; PF00641; zf-RanBP; 1.
 CC SMART; SM00184; RING; 1.
 CC PROSITE; PS01358; ZF_RANBP2_1; 1.
 CC PROSITE; PS01359; ZF_RANBP2_2; 1.
 CC PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
 CC PROSITE; PS00089; ZF_RING_2; 1.
 CC Nuclear protein; Ligase; Ub1 conjugation pathway; Zinc; Zinc-finger;
 KW Metal-binding.
 FT DOMAIN 27 107
 FT DOMAIN 179 185
 FT DOMAIN 190 202
 FT DOMAIN 210 304
 FT DOMAIN 210 215
 FT DOMAIN 242 331
 FT DOMAIN 243 301
 FT ZN_FING 299 328
 FT ZN_FING 438 479
 FT DOMAIN 466 473
 FT
 SQ SEQUENCE 491 AA; 55279 MW; 641E033D5C1DEC39 CRC64;
 Query Match 94.6%; Score 1035.5; DB 1; Length 491;
 Best Local Similarity 94.5%; Pred. No. 8.7e-73;
 Matches 205; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGA VTTTQIPASQETLVPRKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
 DB 1 MCNTNMSVPTDGA VTTTQIPASQETLVPRKPLLLKLLKSVGAQKDTYTMKEVLFYLQY 60

QY 60 IMTKRLYDEKQOHVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDSGT 119
 DB 61 IMTKRLYDEKQOHVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDSGT 120

QY 120 SVSNRCHLEGGSDQKDLVQELQEPSSSHLVSRPSTSSRRRAISSEENSDLSGERQ 179
 DB 121 SVSNRCHLEGGSDQKDLVQELQEPSSSHLVSRPSTSSRRRAISSEENSDLSGERQ 180

QY 180 RKRHKSISLSPDESALCVIREICERSSSSESTG 216
 DB 181 RKRHKSISLSPDESALCVIREICERSSSSESTG 217

RESULT 3
 MD2_CANFA STANDARD; PRT; 487 AA.

AC P56950; O95KN5;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
 DE Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (cdm2).
 GN MDM2.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OC NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE OF 1-484 FROM N.A.
 RX MEDLINE=20218866; PubMed=10754200;
 RA Nasir L., Burr P.D., McFarlane S.T., Gault E., Thompson H.,

RA Argyle D.J.;
RT "Cloning, sequence analysis and expression of the cDNAs encoding the
RT canine and equine homologues of the mouse double minute 2 (mdm2)
RT proto-oncogene.";
RN Cancer Lett. 152:9-13 (2000).
RL [2]
RN SEQUENCE FROM N.A. (ISOFORMS MDM2 AND MDM2-ALPHA).
RP MEDLINE=2006171; PubMed=10597303;
RA Veidhoen N., Mercalfe S., Milner J.;
RT "A novel exon within the mdm2 gene modulates translation initiation in
RT vitro and disrupts the p53-binding domain of mdm2 protein.";
RL Oncogene 18:7026-7033 (1999).
CC -!- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
CC APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
CC FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
CC TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
CC TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS (BY SIMILARITY).
CC -!- COPACITOR: Zinc is required for ubiquitin ligase E3 activity (By
CC similarity).
CC -!- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
CC SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOBLASTOMA PROTEIN
CC (RB). E1A-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION
CC FACTOR (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. Expressed
CC predominantly in the nucleoplasm (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Mdm2;
CC IsoId=p56950-1; Sequence=Displayed;
CC Name=Mdm2-alpha;
CC IsoId=p56950-2; Sequence=VSP 003206;
CC -!- TISSUE SPECIFICITY: Isoform Mdm2-alpha is present in lymphoid and
CC testicular tissues.
CC -!- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
CC G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
CC REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
CC INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
CC FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
CC ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
CC PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
CC ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
CC ITSELF (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the MDM2 / MDM4 family.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -!- SIMILARITY: Contains 1 SWI2 domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF100705; AAF67833.1; -;
DR EMBL; AF322416; AAG42840.1; -;
DR HSSP; Q9UNT8; 1YCR.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf RangDP.
DR InterPro; IPR001841; Znf_rang.
DR Pfam; PF02201; SWIB; 1.
DR Pfam; PF06641; Zf-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01359; ZF_RANBP2_1; 1.
DR PROSITE; PS01359; ZF_RANBP2_2; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE NEG.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Nuclear protein; Ligase; Ub1 conjugation pathway; Zinc; Zinc-finger;
KW Metal-binding; Alternative splicing.
FT DOMAIN 27 107 SWIB.
FT DOMAIN 179 185 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 190 202 NUCLEAR EXPORT SEQUENCE.

FT DOMAIN 210 304 ARF BINDING.
FT DOMAIN 210 215 POLY-SER.
FT DOMAIN 242 331 REGION II.
FT DOMAIN 243 301 ASP/GLU-RICH (ACIDIC).
FT ZN_FING 299 328 RANBP2-TYPE.
FT ZN_FING 434 475 RING-TYPE.
FT DOMAIN 462 469 NUCLEOLAR LOCALIZATION SIGNAL
FT VARSPLIC 1 61 Missing (in isoform Mdm2-alpha).
FT CONFLICT 11 11 /FTId=VSP 003206.
FT CONFLICT 238 239 G -> D (IN REF. 2).
SQ SEQUENCE 487 AA; 54696 MW; 60CDB470A32A8E69 CRC64;

Query Match 91.7%; Score 1004.5; DB 1; Length 487;
Best Local Similarity 92.2%; Pred. No. 2.1e-70;
Matches 200; Conservative 6; Mismatches 10; Indels 1; Gaps 1;

QY 1 MONTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTMKEVLVYL-QY 59
Db 1 MONTNMSVSTGGAVTSTQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTMKEVIFYLGQY 60

QY 60 IMTKLYDEKQOHIVYCSNDLLGLDFGVPSFVKHRRKIYTYRNLVVNOQESSDSGT 119
Db 61 IMTKLYDEKQOHIVYCSNDLLGLDFGVPSFVKHRRKIYTYRNLVVNOHESDSGT 120

QY 120 SVSENRCHEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISEENSDLSGRQ 179
Db 121 SVSENSCHREGGSDQKDPVQELQEKPSDDLISRPSTSSRRRTISEEHADDPGRQ 180

QY 180 RKRHKSDSISLFSDESALCVIREICCRSSSSESTG 216
Db 181 RKRHKSDSISLFSDESALCVIREICCRSSSSESTG 217

RESULT 4
MDM2_MESAU STANDARD; PRT; 466 AA.
ID MDM2_MESAU
AC Q60524;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
DE Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (Fragment).
GN MDM2.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
[1]
RN SEQUENCE FROM N.A.
RP TISSUE=Pancreas;
RX MEDLINE=95300112; PubMed=7780969;
RA Chang K.W., Laconi S., Mangold K.A., Hubchak S., Scarpelli D.G.;
RT "Multiple genetic alterations in hamster pancreatic ductal
RT adenocarcinomas.";
RL Cancer Res. 55:2560-2568 (1995).
CC -!- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
CC APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
CC FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
CC TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
CC TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS (BY SIMILARITY).
CC -!- COPACITOR: Zinc is required for ubiquitin ligase E3 activity (By
CC similarity).
CC -!- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
CC SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOBLASTOMA PROTEIN
CC (RB). E1A-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION
CC FACTOR (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. Expressed
CC predominantly in the nucleoplasm (By similarity).
CC -!- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
CC G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.

REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE CA-TYPE ZINC FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND ITSELF (BY SIMILARITY).

-1- DISEASE: The gene for this protein is overexpressed in some tumors.

-1- SIMILARITY: Belongs to the MDM2 / MDM4 family.

-1- SIMILARITY: Contains 1 RING-type zinc finger.

-1- SIMILARITY: Contains 1 RANBP2-type zinc finger.

-1- SIMILARITY: Contains 1 SWIB domain.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U10982; AAC52425.1; --
DR HSP; Q9UMT3; 1YCR.
DR InterPro; IPR001121; SWIB.
DR InterPro; IPR001876; Znf_RanGDP.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF02201; SWIB; 1.
DR Pfam; PF00641; zf-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS00199; ZF_RANBP2_2; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Nuclear protein; ligase; Ub1 conjugation pathway; Proto-oncogene;
KW Zinc; Zinc-finger; Metal-binding.
FT NON_TER 1 98
FT DOMAIN 19 98 SWIB
FT DOMAIN 169 175 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 176 188 NUCLEAR EXPORT SEQUENCE.
FT DOMAIN 196 201 POLY-SER.
FT DOMAIN 196 290 ARF BINDING.
FT DOMAIN 228 317 REGION II.
FT DOMAIN 209 287 ASP/GLU-RICH (ACIDIC).
FT ZN_FING 285 314 RANBP2-TYPE.
FT ZN_FING 419 460 RING-TYPE.
FT DOMAIN 447 454 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT NON_TER 466 466
FT SEQUENCE 466 AA; 52390 MW; 78A3042163C5F939 CRC64;
Query Match 76.9%; Score 842.5; DB 1; Length 466;
Best Local Similarity 83.5%; Pred. No. 6.6e-58;
Matches 172; Conservative 12; Mismatches 17; Indels 5; Gaps 2;
QY 10 TDGAVTTSQIPASEQETLVREPKLLKLSVGAQKDTYTMKEVLVYQVYIMTKRLYDEK 69
Db 2 TDGAGTSQIPASEQETLVREPKFLKLSVGAQKDTYTMKEIILISQVYIMTKRLYDEK 61
QY 70 QQHIVYCSNLLGLDFGVPSFVKRHKIYTYMIRNLVYNQESDSTGTSVSENCHLE 129
Db 62 QQHIVYCSNLLGLDFGVPSFVKRHKIYTYMIRNLVYNVVSQQTLSQSTSVSESRCPQE 121
QY 130 GGSDDKDLVQLQEEKSSHLVSRPSTSSRRRAISSTENSDLSGERQKHKSDSY 189
Db 122 GGEQKDPVQPEEK-SSSDSVSRPSTSSRRRAISSTENADELPDRQKHKR-----S 176
QY 190 LSFDESIALCVIRBICCCSSSEST 215
Db 177 LSFDESIALCVLRHICCCSSSEST 202

RESULT 5

MDM2 MOUSE
ID MDM2 MOUSE STANDARD; PRT; 489 AA.
AC P23804; Q61040; Q64330;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein).
GN MDM2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
[1]
RP SEQUENCE FROM N.A. (ISOFORM MDM2-P90).
RX MEDLINE=91224107; PubMed=2026149;
RA Fakharzadeh S.S., Trusko S.P., George D.L.;
RT "Tumorigenic potential associated with enhanced expression of a gene that is amplified in a mouse tumor cell line."
RL EMBO J. 10:1565-1569(1991).
[2]
RP SEQUENCE FROM N.A. (ISOFORM MDM2-P90).
RX STRAIN=129/Sv;
RA Jones S.N., Ansari-Lari M.A., Hancock A.R., Jones W.J., Gibbs R.A.,
RA Donenower L.A., Bradley A.;
RT "Genomic organization of the mouse double minute 2 gene."
RL Gene 175:209-213(1996).
[3]
RP SEQUENCE FROM N.A. (ISOFORM MDM2-P90).
RX STRAIN=129/Sv;
RA de Oca Luna R.M., Tabor A.D., Eberspaecher H., Hulboy D.L.,
RA Worth L.L., Colman M.S., Finlay C.A., Lozano G.;
RT "The organization and expression of the mdm2 gene."
RL Genomics 33:352-357(1996).
[4]
RP SEQUENCE FROM N.A. (ISOFORMS MDM2-P90 AND MDM2-P76).
RX MEDLINE=99175199; PubMed=10075719;
RA Saucedo L.J., Myers C.D., Perry M.E.;
RT "Multiple murine double minute gene 2 (MDM2) proteins are induced by ultraviolet light."
RL J. Biol. Chem. 274:8161-8168(1999).
[5]
RP NUCLEAR LOCALIZATION SIGNAL.
RX MEDLINE=20180080; PubMed=10713175;
RA Weber J.D., Kuo M.-L., Bothner B., DiGiannmarino E.L., Kriwacki R.W.,
RA Roussel M.F., Sherr C.J.;
RT "Cooperative signals governing ARF-mdm2 interaction and nuclear localization of the complex."
RL Mol. Cell. Biol. 20:2517-2528(2000).
[6]
RP PHOSPHORYLATION BY ATM.
RX MEDLINE=20079591; PubMed=10611322;
RA Khosravi R., Maya R., Gottlieb T., Oren M., Shiloh Y., Shkedy D.;
RT "Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation in response to DNA damage."
RL Proc. Natl. Acad. Sci. U.S.A. 96:14973-14977(1999).
-1- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN. FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2, TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS.
-1- COFACTOR: Zinc is required for ubiquitin ligase E3 activity.
-1- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOBLASTOMA PROTEIN (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION FACTOR.
-1- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC. EXPRESSED PREDOMINANTLY IN THE NUCLEOLUS. INTERACTION WITH ARF(P14) RESULTS IN THE LOCALIZATION OF BOTH PROTEINS TO THE NUCLEOLUS. THE NUCLEOLAR LOCALIZATION SIGNALS IN BOTH ARF(P14) AND MDM2 MAY BE NECESSARY TO ALLOW EFFICIENT NUCLEOLAR LOCALIZATION OF BOTH

CC CC PROTEINS

CC CC -1- ALTERNATIVE PRODUCTS:

CC CC Event=Alternative splicing; Named isoforms=2;

CC CC Name=Mdm2-p90;

CC CC IsoId=P223804-1; Sequence=Displayed;

CC CC Note=Isoform Mdm2-p76 can also be produced by alternative

CC CC initiation at Met-50 of isoform Mdm2-p90, but is produced more

CC CC efficiently by alternative splicing;

CC CC Name=Mdm2-p76;

CC CC IsoId=P223804-2; Sequence=VSP_003215;

CC CC Note=Does not bind to p53;

CC CC Event=Alternative initiation;

CC CC Comment=2 isoforms, Mdm2-p90 (shown here) and Mdm2-p76, are

CC CC produced by alternative initiation at Met-1 and Met-50. Isoform

CC CC Mdm2-p76 is produced more efficiently by alternative splicing;

CC CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED AT LOW-LEVEL THROUGHOUT

CC CC EMBRYO DEVELOPMENT AND IN ADULT TISSUES. MDM2-P90 IS MUCH MORE

CC CC ABUNDANT THAN MDM2-P76 IN TESTIS, BRAIN, HEART, AND KIDNEY, BUT IN

CC CC THE THYMUS, SPLEEN, AND INTESTINE, THE LEVELS OF THE MDM2 PROTEINS

CC CC ARE ROUGHLY EQUIVALENT.

CC CC -1- INDUCTION: BY UV LIGHT.

CC CC -1- DOMAIN: REGION 1 IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS

CC CC G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.

CC CC REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR

CC CC INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE CA-TYPE ZINC

CC CC FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF

CC CC ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS

CC CC PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS

CC CC ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND

CC CC ITSELF.

CC CC -1- PTM: PHOSPHORYLATED IN RESPONSE TO IONIZING RADIATION IN AN ATM-

CC CC DEPENDENT MANNER.

CC CC -1- DISEASE: The gene for this protein is amplified in a mouse tumor

CC CC cell line.

CC CC -1- SIMILARITY: Belongs to the MDM2 / MDM4 family.

CC CC -1- SIMILARITY: Contains 1 RING-type zinc finger.

CC CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.

CC CC -1- SIMILARITY: Contains 1 SWIB domain.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC CC the European Bioinformatics Institute. There are no restrictions on its

CC CC use by non-profit institutions as long as its content is in no way

CC CC modified and this statement is not removed. Usage by and for commercial

CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC CC or send an email to license@isb-sib.ch).

CC CC -----

CC CC EMBL; X58876; CAA41684.1; -

CC CC EMBL; U40145; AAB91167.1; -

CC CC EMBL; U47944; AAB09030.1; JOINED.

CC CC EMBL; U47935; AAB09030.1; JOINED.

CC CC EMBL; U47936; AAB09030.1; JOINED.

CC CC EMBL; U47937; AAB09030.1; JOINED.

CC CC EMBL; U47938; AAB09030.1; JOINED.

CC CC EMBL; U47939; AAB09030.1; JOINED.

CC CC EMBL; U47940; AAB09030.1; JOINED.

CC CC EMBL; U47941; AAB09030.1; JOINED.

CC CC EMBL; U47942; AAB09030.1; JOINED.

CC CC EMBL; U47943; AAB09030.1; JOINED.

CC CC EMBL; U47934; AAB09031.1; -

CC CC PIR; S15349; S15349.

CC CC HSP; Q9UMT8; 1YCR.

CC CC MGD; MGI:96952; Mdm2

CC CC GO; GO:0005634; C:nucleus; IDA.

CC CC GO; GO:0005515; F:protein binding; IPI.

CC CC GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.

CC CC GO; GO:0030163; F:protein catabolism; IDA.

CC CC GO; GO:0016567; P:protein ubiquitination; IDA.

CC CC GO; GO:0007089; P:start control point of mitotic cell cycle; IDA.

CC CC InterPro; IPR003121; SWIB

CC CC InterPro; IPR001876; Znf_RangFP.

CC CC InterPro; IPR001841; Znf_Rang.

CC CC Pfam; PF02201; SWIB; 1.

DR Pfam; PF00641; zf-RanBP; 1.

DR SMART; SMO0184; RING; 1.

DR PROSITE; PS01358; ZF_RANBP2_1; 1.

DR PROSITE; PS50199; ZF_RANBP2_2; 1.

DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.

DR PROSITE; PS00089; ZF_RING_2; 1.

DR PROSITE; PS00089; ZF_RING_2; 1.

KW Nuclear protein; Ligase; Ubl conjugation pathway; Proto-oncogene;

KW Alternative splicing; Ligase; Ubl conjugation pathway; Proto-oncogene;

KW Metal-binding; Phosphorylation.

FT CHAIN 1 489

FT CHAIN 50 489

FT CHAIN 50 489

FT INIT MET 50 107

FT DOMAIN 27 107

FT DOMAIN 176 182

FT DOMAIN 183 195

FT DOMAIN 203 213

FT DOMAIN 208 302

FT DOMAIN 240 329

FT DOMAIN 221 299

FT ZN_FING 237 326

FT ZN_FING 436 477

FT ZN_FING 464 471

FT MOD_RES 1 1

FT VARSP 1 49

FT CONFLICT 203 203

FT CONFLICT 419 419

FT CONFLICT 486 486

FT CONFLICT 489 AA; 54543 MW; 4ABF489E92038DF4 CRC64;

FT SEQUENCE 489 AA; 54543 MW; 4ABF489E92038DF4 CRC64;

Query Match 73.5%; Score 805; DB 1; Length 489;

Best Local Similarity 76.9%; Pred. NO. 5.5e-55;

Matches 166; Conservative 17; Mismatches 25; Indels 8; Gaps 3;

QY 1 MCNTNMSVPTGCAVTTSOIPASEQETLVPRPKPLLLKLLKSVGAQNDTYTKEVLYL-QY 59

Db 1 MCNTNMSVSTEGAASTQIPASEQETLVPRPKPLLLKLLKSVGAQNDTYTKEIIFYGY 60

QY 60 IMTKELYDEKQHIYVCNDLLGLDVGVPFSVKEHRIYMIYRNLVVNVQSSDSGT 119

Db 61 IMTKELYDEKQHIYVCNDLLGLDVGVPFSVKEHRIYMIYRNLVVNVQSSDSGT 117

QY 120 SVSENRCHLEGSDQKDLVQLQEKPSHIVSPSTSSRRATSETEENSDLSGERQ 179

Db 118 SLSESRROEGGSDLKDFLQAPPEKPSDDLISLSTSSRRRSISETEENIDELPGERH 177

QY 180 RKRHKSDSISLSDSLALCVIREICCRSSSSSEST 215

Db 178 RKRR-----SLSFDPISLGLCELRENCSCGSSSSSSS 209

RESULT 6

ID_MDM2_XENLA STANDARD; PRT; 473 AA.

AC P56273;

DT 15-JUL-1998 (Rel. 36, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein

DE Mdm2) (Double minute 2 protein) (Xdm2).

GN MDM2.

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

OC Xenopodinae; Xenopus.

CC NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-9728206; PubMed-9136986;

RA Marechal V., Elenbaas B., Taneyhill L., Piette J., Mechali M.,

RA Nicolas J.-C., Levine A.J., Moreau J.;
RT "Conservation of structural domains and biochemical activities of the
RT MDM2 protein from *Xenopus laevis*;"
RN Oncogene 14:1427-1433(1997).
RL (2)
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 13-119 IN COMPLEX WITH P53.
RX MEDLINE=97081050; PubMed=8875929;
RA Kussie P.H., Gorina S., Marechal V., Elenbaas B., Moreau J.,
RA Levine A.J., Pavletich N.P.;
RT "Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor
RT transactivation domain;"
RL Science 274:948-953(1996).
CC -!- FUNCTION: May bind p53 protein and may function as a ubiquitin
CC ligase E3.
CC -!- COFACTOR: Zinc is required for ubiquitin ligase E3 activity (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- DEVELOPMENTAL STAGE: EXPRESSION INCREASES FROM OOCYTE STAGE I/II
CC TO REACH ITS MAXIMUM IN OOCYTE STAGE V/VI IN UNFERTILIZED EGGS,
CC AND THEN PROGRESSIVELY DECREASES TO BECOME UNDETECTABLE AT THE
CC GASTRULA STAGE.
CC -!- SIMILARITY: Belongs to the MDM2 / MDM4 family.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -!- SIMILARITY: Contains 1 SWIB domain.
DR PDB; 1YCQ; 19-NOV-97.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf_RanGDP.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF02201; SWIB; 1.
DR Pfam; PF00641; Zf-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00547; Znf_RBZ; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS00199; ZF_RANBP2_2; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS00089; ZF_RING_2; 1.
DR Nuclear protein; Ligase; Ub1 conjugation pathway; Zinc; Zinc-finger;
KW Metal-binding; 3D-structure.
FT DOMAIN 23 103 SWIB.
FT DOMAIN 173 179 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 230 322 REGION II.
FT DOMAIN 216 292 ASP/GLU-RICH (ACIDIC).
FT ZN_FING 290 319 RANBP2-TYPE.
FT ZN_FING 420 461 RING-TYPE.
FT STRAND 24 26
FT HELIX 28 36
FT TURN 37 38
FT STRAND 44 44
FT HELIX 46 60
FT TURN 61 61
FT STRAND 63 63
FT TURN 68 69
FT STRAND 70 72
FT TURN 74 75
FT HELIX 77 82
FT TURN 83 83
FT STRAND 86 88
FT TURN 89 90
FT HELIX 92 100
FT TURN 101 102
FT STRAND 103 105
SQ SEQUENCE 473 AA; 53464 MW; 7DA668DE8B3BEE01 CRC64;
Query Match 45.4%; Score 497; DB 1; Length 473;
Best Local Similarity 54.0%; Pred. No. 3.3e-31;
Matches 114; Conservative 36; Mismatches 49; Indels 12; Gaps 7;
QY 10 TDGAVTTSOIPASEGETVVRPKPLLLKLLKSGAKDVTYMKELVYL-QYIMTKRLYDE 68
DB 6 TTNENNNHISTSDGKLVQPTPLLSLLKSGAKQFTMKELVYHUGQYIMAKQLYDE 65
QY 69 KQCHIVYCSNDLLGLDPLFGVPVSVKHEKRIYTMVYRNVLVWVQNQSSSDSGTSVSENRC 128

Db 66 KQCHIVYCSNDPLGELFGVQVSVKPRRLYAMISRNLSVANKESSE---DIFGNVCCF 122
QY 129 -EGSDOKDLVQLQEE--KPSSSHVSRPSTSSRRRAISSETEE--NSDELSGRQRKHK 184
DB 123 PDKOSSQKQLQELPDKLIAPSD---SKPCNUSQRKSSNETETSSVDHPASQQRKHK 179
QY 185 SDGISLSFDESALCVIREICCRSSSESST 215
DB 180 SDGSLTFDESLSMWVIGLRCRDR-NSSEST 209

RESULT 7
MDM2 BRARE STANDARD; PRT; 445 AA.
AC 042354;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
DE Mdm2) (Double minute 2 protein).
GN MDM2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RX SEQUENCE FROM N.A.
RA Neal H., Pieter J.;
RT "Partial cDNA nucleotide sequence of the zebrafish homolog of Mdm2.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May bind p53 protein and may function as a ubiquitin
CC ligase E3.
CC -!- COFACTOR: Zinc is required for ubiquitin ligase E3 activity (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the MDM2 / MDM4 family.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -!- SIMILARITY: Contains 1 SWIB domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF010255; AAB64176.1; -.
DR HSPF; P56273; LYCQ.
DR ZFIN; ZDB-GENE-990415-153; mdm2.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf_RanGDP.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF02201; SWIB; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS00199; ZF_RANBP2_2; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Nuclear protein; Ligase; Ub1 conjugation pathway; Zinc; Zinc-finger;
KW Metal-binding.
FT DOMAIN 20 100 SWIB.
FT DOMAIN 141 145 POLY-ARG.
FT DOMAIN 160 166 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 171 183 NUCLEAR EXPORT SEQUENCE.
FT DOMAIN 190 279 ARF BINDING.
FT DOMAIN 222 306 REGION II.
FT DOMAIN 210 276 ASP/GLU-RICH (ACIDIC).
FT ZN_FING 274 303 RANBP2-TYPE.
FT ZN_FING 392 433 RING-TYPE.

```
FT DOMAIN 420 427 NUCLEOLAR LOCALIZATION SIGNAL
FT (POTENTIAL)
SQ SEQUENCE 445 AA; 49949 MW; 6FA8175A8A8E6261 CRC64;

Query Match 37.3%; Score 408.5; DB 1; Length 445;
Best Local Similarity 45.9%; Pred. No. 2.1e-24;
Matches 96; Conservative 29; Mismatches 65; Indels 19; Gaps 4;

QY 10 TDGAVTTQIPASGETLVRPKPLLLKLLKSVGAQKTYTMKEVLYL-QYIMTKRYLYDE 68
D5 3 TESCLSSQISKEVDNEKLVKPVQKSLLEDAGADKDVFTKWEVFLGYIKTIYSKEDYDK 62
QY 69 KQOHIVYCNLDLGLFGVPSPSVKHEHRIYTMIRNLVNVNQESSDSGTSVSENKCHL 128
D5 63 QOOHIVHCGEDPLGAVLGVKFSVKEPRALPALINRLNLTVPKNEPQ---STSEPRSQS 119
QY 129 EGGSDQKDLVELDEKPPSSHLVSRPSTSS---RRRAISETEENSDELSEGERKHKHS 185
D5 120 E-----PDRGPGDTSDSRSSTSQQRRRRSSDPSESSAEDSRERRKHKHS 167
QY 186 DSISLSPDESIALCVIREICCRSSSSSES 214
D5 168 DPSLITFDLSLWCVIGLHREGNSSES 196

RESULT 8
MDM4 MOUSE
ID MDM4 MOUSE STANDARD; PRT; 489 AA.
AC O35618;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Mdm4 protein (p53-binding protein Mdm4) (Mdm2-like p53-binding
DE protein) (Mdmx protein) (Double minute 4 protein).
GN MDM4 OR MDMX.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RC TISSUE=Embryo;
RX MEDLINE=97050840; PubMed=8995579;
RA Shvarts A., Steegenga W.T., Riteco N., Van Laar T., Dekker P.,
RA Bazuine M., Van Ham R.C.A., Van der Houven van Oordt W., Hateboer G.,
RA Van der Eb A.J., Jochensen A.G.;
RT "MDMX: a novel p53-binding protein with some functional properties of
RT MDM2."
RL EMBO J. 15:5349-5357(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM LONG).
RA Jochensen A.G., Shvarts A., Steegenga W.T., Riteco N., Van Laar T.,
RA Dekker P., Bazuine M., Van Ham R.C.A., van Oordt W., Hateboer G.,
RA der Eb A.J.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).
RX MEDLINE=99175216; PubMed=10075736;
RA Rallapalli R., Strachan G., Cho B., Mercer W.E., Hall D.J.;
RT "A novel MDMX transcript expressed in a variety of transformed cell
RT lines encodes a truncated protein with potent p53 repressive
RT activity."
RL J. Biol. Chem. 274:8299-8308(1999).
CC -1- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
CC APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
CC INHIBITS DEGRADATION OF MDM2. CAN REVERSE MDM2-TARGETED
CC DEGRADATION OF P53 WHILE MAINTAINING SUPPRESSION OF P53
CC TRANSACTIVATION AND APOPTOTIC FUNCTIONS. THE SHORT ISOFORM IS A
CC MORE POTENT INHIBITOR OF P53 ACTIVITY THAN THE LONG ISOFORM.
CC -1- SUBUNIT: BINDS TO P53, P73 AND MDM2.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
```


RC TISSUE=Liver;
RX MEDLINE=66148479; PubMed=3005973;
RA Maguire D.J.; Day A.R., Borthwick I.A., Srivastava G., Wigley P.L.,
RA May B.K., Elliott W.H.;
RT "Nucleotide sequence of the chicken 5-aminolevulinate synthase gene.";
RL Nucleic Acids Res. 14:1379-1391(1986).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=85257679; PubMed=3839458;
RA Borthwick I.A., Srivastava G., Day A.R., Pirola B.A., Snowell M.A.,
RA May B.K., Elliott W.H.;
RT "Complete nucleotide sequence of hepatic 5-aminolevulinate synthase
precursor.";
RL Eur. J. Biochem. 150:481-484(1985).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=8912863; PubMed=2915978;
RA Riddle R.D., Yamamoto M., Engel J.D.;
RT "Expression of delta-aminolevulinate synthase in avian cells:
separate genes encode erythroid-specific and nonspecific isozymes.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:792-796(1989).
CC -1- CATALYTIC ACTIVITY: Succinyl-CoA + glycine = 5-aminolevulinate +
CoA + CO(2).
CC -1- COFACTOR: Pyridoxal phosphate.
CC -1- PATHWAY: Heme biosynthesis; first (rate-limiting) step.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- MISCELLANEOUS: THERE ARE TWO DELTA-ALA SYNTHETASE IN VERTEBRATES:
AN ERYTHROID-SPECIFIC FORM AND ONE (HOUSEKEEPING) WHICH IS
EXPRESSED IN ALL TISSUES.
CC -1- SIMILARITY: Belongs to class-II of pyridoxal-phosphate-dependent
aminotransferases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC -----
CC EVBL; X02827; CAA26595.1; --
CC EVBL; X03517; CAA27223.1; --
CC EVBL; X03627; CAA27223.1; JOINED.
CC PIR; A23538; SYCHAL.
CC HGSP; P12998; 1RSO.
CC InterPro; IPR003408; Ala synthase.
CC InterPro; IPR004839; Aminotrans I/II.
CC InterPro; IPR001917; Aminotrans II.
CC Pfam; PF02490; Ala synthase; 1.
CC Pfam; PF00155; aminotran_1_2; 1.
CC DR PROSITE; PS00599; AA TRANSFER CLASS 2; 1.
CC Heme biosynthesis; Transferase; Acyltransferase; Mitochondrion;
KW Transit peptide; Pyridoxal phosphate; Multigene family.
FT TRANSIT 1 56 MITOCHONDRION
FT CHAIN 57 635 5-AMINOLEVULINIC ACID SYNTHASE,
NONSPECIFIC.
FT BINDING 440 440 PYRIDOXAL PHOSPHATE (PROBABLE).
FT CONFLICT 53 53 S -> A (IN REF. 2 AND 3).
SQ SEQUENCE 635 AA; 69948 MW; EEE708B85C1539BA CRC64;
Query Match 8.8%; Score 96.5; DB 1; Length 635;
Best Local Similarity 20.9%; Pred. No. 4;
Matches 65; Conservative 37; Mismatches 84; Indels 125; Gaps 14;
Cc
Cc 8 VFTDGVTTSGIPASEQETLVRP-KPLLLK-----LLKSVGA----- 43
Cc
Cc 142 IPTSVVNTAEAGEQGLKKFKDMLKQRPESVSHLLQDNLPKSVSFQYDQPFKK 201
Cc
Cc 44 -----QKD-----TYTMKEVLP-----YLOYIMTKELYDEKQHYVCSNDLLG----- 82
Cc
Cc 202 IDEKKDHTYRVFTVNRKAQIPPMADDYSDSLITKK-----EVSVCNSDYLGMGRHP 255

QY 83 -----DLQVPSFSV---KE-----HRKIYTMIRNLVVNQ--- 111
Db 256 RVCGAVMTLKHQAGAGCTNIGTSKFHVDLEKELADLHGKDAALLFSSCFVANDSTL 315
QY 112 -----QSSDSQTSVS-----ENRC--HLEGSDDQDLVQLQESKPSSSHLYS 153
Db 316 FTAKMLPGCEIYSDSGNHASMIQGINRSRVPKHFIRHNDVNLHRLKSLKSDPTPKIVA 375
QY 154 RPSISSRRRAISETEENSDELS-----GSRQRKRHKSIS 189
Db 376 FETVHMDGAVCPLEELCDVAHEGATFTFDEVHANGVLYGARGGGIGDRDGMHMDIIS 435
QY 190 LSFDESALCV 200
Db 436 GTLGKAFV-CV 445
RESULT 11
MYSC CHICK
ID MYSC CHICK STANDARD; PRT; 1102 AA.
AC P29616;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain, cardiac muscle isoform (Fragment).
OS Gallus gallus (Chicken).
OC Archosauromorpha; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus
OC NCBI_TaxID=9031;
RN [1]
RP SEQUENCE OF 65-1102 FROM N.A.
RC STRAIN=Broiler; TISSUE=Heart;
RX MEDLINE=92130260; PubMed=1774788;
RA Stewart A.F.R., Camoretti-Mercado B., Perlman D., Gupta M.,
RA Jakovcic S., Zak R.;
RT "Structural and phylogenetic analysis of the chicken ventricular
myosin heavy chain rod.";
RL J. Mol. Evol. 33:357-366(1991).
RN [2]
RP SEQUENCE OF 1-259.
RC TISSUE=Heart;
RX MEDLINE=93039740; PubMed=1418675;
RA Watanabe B.;
RT "Amino-acid sequence of the short subfragment-2 in adult chicken
cardiac muscle myosin.";
RL Biol. Chem. Hoppe-Seyler 373:1045-1054(1992).
CC -1- FUNCTION: Muscle contraction.
CC -1- SUBUNIT: Muscle myosin is a hexameric protein that consists of 2
heavy chain subunits (MHC), 2 alkali light chain subunits (MLC)
and 2 regulatory light chain subunits (MLC-2).
CC -1- SUBCELLULAR LOCATION: Thick filaments of the myofibrils.
CC -1- TISSUE SPECIFICITY: Ventricular muscle and traces levels in the
atrium. Also transient expression in skeletal muscle during fetal
development and regeneration following freeze injury.
CC -1- DOMAIN: The rodlike tail sequence is highly repetitive, showing
cycles of a 28-residue repeat pattern composed of 4 heptapeptides,
characteristic for alpha-helical coiled coils.
CC -1- MISCELLANEOUS: Each myosin heavy chain can be split into 1 light
meromyosin (LMM) and 1 heavy meromyosin (HMM). It can later be
split further into 2 globular subfragments (S1) and 1 rod-shaped
subfragment (S2).
CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

DR EMBL; X59552; CAA42130.1; --
DR HSP; P03437; 1HTM.
DR InterPro; IPR002928; Myosin tail.
DR Pfam; PF01576; Myosin tail; 1.
KW Myosin; Muscle protein; Coiled coil; Thick filament; Actin-binding;
KW ATP-binding; Multigene family.
FT NON_TER 1 1
FT DOMAIN <1 1102 COILED COIL (POTENTIAL).
FT CONFLICT 65 65 A -> T (IN REF. 1).
FT CONFLICT 92 92 V -> A (IN REF. 1).
FT CONFLICT 113 114 CS -> FAL (IN REF. 1).
FT CONFLICT 135 135 H -> Q (IN REF. 1).
FT CONFLICT 185 185 T -> R (IN REF. 1).
SQ SEQUENCE 1102 AA; 128008 MW; 2293668D049825DC CRC64;
Query Match 8.4%; Score 92; DB 1; Length 1102;
Best Local Similarity 21.3%; Pred. No. 17;
Matches 42; Conservative 40; Mismatches 63; Indels 52; Gaps 7;
QY 37 LLKSVGAQDTYMKVFLVLYIMYX-----RLYDEKQHIYVCNDLLGLDFGVSPS 91
DB 5 LLKSAETKEMANNKEFLKLEALPKSEARRKEEKQVSLVQEKNDLL----- 54
QY 92 VKHRIKTYMYRNLYVNVNQESSDSGTSENRCHE-----EGSSDOKDLVQELQREK 145
DB 55 -----LQLQAEQTLDAERCDLLKSKQLQBAKVEITERVEDSE 96
QY 146 PSSSHLVSRPSTSSRR--AISETENSDLS--GERQKRHSKDSLSLSDSLALC 199
DB 97 ENNSL-----TSKKKLEDCSELKDDIDLEITLAKVEKEKHAKENKVNLTETMAYL 151
QY 200 --VIRICERSSSSS 214
DB 152 DENISKLTREKSLQEA 168
RESULT 12
YH85 SCHPO
ID YH85 SCHPO STANDARD; PRT; 1465 AA.
AC QP5N0; Q9U044;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable ATP-dependent permealase C359.05.
GN SPBC359.05.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
CC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Haidago J., Hodgson G.,
RA Holroyd S., Horisby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Ruster S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkart R.G., Aert R., Robben J., Grymopoulos B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritsch C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,

RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe";
RN Nature 415:871-880(2002).
RL Nature 415:871-880(2002).
RN [2]
RP SEQUENCE OF 207-348 FROM N.A., AND SUBCELLULAR LOCATION.
RC STRAIN=968 h90;
RX MEDLINE=20233868; PubMed=10759889;
RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RA Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
fission yeast cells by the use of a GFP-fusion genomic DNA library";
RL Genes Cells 5:169-190(2000).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the ABC transporter family, MRP subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
EMBL; AL356012; CAB91574.1; --
EMBL; AB027821; BAA87125.1; --
DR GeneDB SPombe; SPBC359.05; --
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR001140; ABC_TM_transpt.
DR InterPro; IPR003439; ABC_transporter.
DR Pfam; PF00664; ABC_membrane; 2.
DR Pfam; PF00005; ABC_tran; 2.
DR ProDom; PD000006; ABC_transporter; 2.
DR SMART; SMO0382; AAA; 2.
DR PROSITE; PS50929; ABC_TM1F; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR PROSITE; PS50893; ABC_TRANSPORTER_2; 2.
KW ATP-binding; Transmembrane; Glycoprotein; Transport; Repeat.
FT DOMAIN 1 6 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 7 29 POTENTIAL.
FT DOMAIN 30 43 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 44 63 POTENTIAL.
FT DOMAIN 64 69 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 70 89 POTENTIAL.
FT DOMAIN 90 101 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 102 124 POTENTIAL.
FT DOMAIN 125 235 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 236 255 POTENTIAL.
FT DOMAIN 256 292 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 293 310 POTENTIAL.
FT DOMAIN 311 365 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 366 388 POTENTIAL.
FT DOMAIN 389 392 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 393 415 POTENTIAL.
FT DOMAIN 416 478 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 479 501 POTENTIAL.
FT DOMAIN 502 515 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 516 538 POTENTIAL.
FT DOMAIN 539 900 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 901 923 POTENTIAL.
FT DOMAIN 924 942 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 943 965 POTENTIAL.
FT DOMAIN 966 1031 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1032 1054 POTENTIAL.
FT DOMAIN 1055 1465 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1465 804 ABC_TRANSPORTER_1.
FT DOMAIN 804 804 ABC_TRANSPORTER_2.
FT DOMAIN 1226 1460 ATP (POTENTIAL).
FT NP_BIND 414 521 ATP (POTENTIAL).
FT NP_BIND 1260 1267 ATP (POTENTIAL).
FT CARBOHYD 346 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT 346

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: March 14, 2004, 21:56:30 ; Search time 72.9931 Seconds
(without alignments)
1642.202 Million cell updates/sec

Title: US-10-057-510-4
Perfect score: 1095
Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIREICCRSSSESTG 216

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xih
-DB=frame+ p2n.model -DEV=xih
-DB=Issued Patents NA -QPMT=fastap -SUFFIX=std.rni -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10057510 @CGN 1.1.63 @runat 09032004 162235 26011 -NCPUS=6 -ICPU=3
-NC MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfilesi.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1084.5	99.0	652	4	US-09-510-252-3
2	1084.5	99.0	2372	1	US-07-903-103-1
3	1084.5	99.0	2372	1	US-08-044-619A-1
4	1084.5	99.0	2372	1	US-08-283-911-1
5	1084.5	99.0	2372	1	US-08-245-500A-2
6	1084.5	99.0	2372	1	US-08-390-546-2
7	1084.5	99.0	2372	1	US-08-390-479A-2
8	1084.5	99.0	2372	1	US-08-557-393-2
9	1084.5	99.0	2372	1	US-08-390-516C-2
10	1084.5	99.0	2372	1	US-08-390-517A-2
11	1084.5	99.0	2372	1	US-08-390-515A-2
12	1084.5	99.0	2372	2	US-08-801-718-2

13	1084.5	99.0	2372	3	US-09-073-567-1	Sequence 1, Appli
14	1084.5	99.0	2372	3	US-09-280-805-1	Sequence 1, Appli
15	1084.5	99.0	2372	3	US-09-048-810-1	Sequence 1, Appli
16	1084.5	99.0	2372	4	US-09-170-159A-2	Sequence 2, Appli
17	1084.5	99.0	2372	4	US-09-480-718-43	Sequence 43, Appli
18	940.5	85.9	729	4	US-09-603-052-3	Sequence 3, Appli
19	804.5	73.5	1710	1	US-07-903-103-3	Sequence 3, Appli
20	804.5	73.5	1710	1	US-08-044-619A-3	Sequence 3, Appli
21	804.5	73.5	1710	1	US-08-283-911-3	Sequence 3, Appli
22	804.5	73.5	1710	1	US-08-245-500A-4	Sequence 4, Appli
23	804.5	73.5	1710	1	US-08-390-546-4	Sequence 4, Appli
24	804.5	73.5	1710	1	US-08-390-479A-4	Sequence 4, Appli
25	804.5	73.5	1710	1	US-08-557-393-4	Sequence 4, Appli
26	804.5	73.5	1710	1	US-08-390-516C-4	Sequence 4, Appli
27	804.5	73.5	1710	1	US-08-390-517A-4	Sequence 4, Appli
28	804.5	73.5	1710	2	US-08-390-515A-4	Sequence 4, Appli
29	804.5	73.5	1710	2	US-08-801-718-4	Sequence 4, Appli
30	804.5	73.5	1710	3	US-09-073-567-12	Sequence 12, Appli
31	804.5	73.5	1710	4	US-09-170-159A-4	Sequence 4, Appli
32	804.5	73.5	1710	4	US-09-480-718-45	Sequence 45, Appli
33	378	34.5	309	4	US-09-167-322-9	Sequence 9, Appli
34	283.5	25.9	966	4	US-09-167-322-7	Sequence 7, Appli
35	254	23.2	2192	3	US-09-289-267-1	Sequence 1, Appli
36	214	19.5	199	4	US-09-200-355-1	Sequence 1, Appli
37	214	19.5	199	4	US-09-200-355-2	Sequence 2, Appli
38	157	14.3	891	4	US-09-167-322-5	Sequence 5, Appli
39	153	14.0	657	4	US-09-167-322-6	Sequence 6, Appli
40	147	13.4	399	4	US-09-167-322-8	Sequence 8, Appli
41	125	11.4	73	3	US-09-073-567-49	Sequence 49, Appli
42	119	10.9	500	3	US-09-280-805-2	Sequence 2, Appli
43	119	10.9	500	3	US-09-048-810-2	Sequence 2, Appli
44	100	9.1	10640	4	US-09-417-485D-5	Sequence 5, Appli
45	93	8.5	3114	3	US-08-946-026-12	Sequence 12, Appli

ALIGNMENTS

RESULT 1
US-09-510-252-3
; Sequence 3, Application US/09510252
; Patent No. 6372490
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDM INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT APPLICATION NUMBER: US/09/510,252
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USSN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USSN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 652
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-510-252-3

Alignment Scores:
Pred. No.: 3,198-116 Length: 652
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 4 Gaps: 1

US-10-057-510-4 (1-216) x US-09-510-252-3 (1-652)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerglnIlePro 20

DB: 1 Gaps: 1
US-10-057-510-4 (1-216) x US-08-245-500A-2 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTGTAAACACCTCACAGATTCCA 371
QY 21 AlasrGluGlnGluThrValArgProLysProLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 372 GCTTCGGAACAAGAGACCTGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrValMetLysGluValLeuPheValThrLeuLeuLeuLeuLeu 59
Db 432 GTTGTGCAACAAAAGACACTTACTATGAAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuValProSerPheSerValLysGlnHisLeuValThrCysSerAsnAsp 79
Db 492 ATTATGACTAAACGATTATATGATGAGAAGCAACATATTGTTATTTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGlnHisArgLysIleValThr 99
Db 552 CTTCTAGGAGATTGTTGGCGTCCAGCTTCTGTGAAGAGAGACAGGAAATATAT 611
QY 100 ThrMetIleValArgLeuValValAsnGlnGlnLysSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGGAACCTTGTAGTCAATCAGCAGGAATCATCGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTAGTGAGAACAGGTTGACCTTTGAAGAGGTTGATCAAAAGGACCTTGTACAA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GACCTTCAGGAAGAGAAACCTTCTTCATCTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgAlaIleSerGluThrGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGAGAGACCAATTAGTGAGACAGAGAAATTCAGATGAATTCGTGTAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGlnSerLeuAlaLeuCys 199
Db 852 AGAAACGCCACAACTGATGATATTTCCCTTCTTTGATGAAGGCTTGGCTCTGTGT 911
QY 200 ValIleArgGluLeuCysGlyArgSerSerSerSerSerSerSerSerSerSerSerSer 216
Db 912 GTAATAAGGAGATATGTTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 962

RESULT 6

US-08-390-546-2
; Sequence 2, Application US/08390546
; Patent No. 560604
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOSELSSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/390,546
; FILING DATE: 07-APR-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107,42798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BMB UT
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2372 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL LINE: CaCo-2
; POSITION IN GENOME:
; MAP POSITION: 12q12-14
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 312..1784
; US-08-390-546-2
Alignment Scores:
Pred. No.: 2,12e-115 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
Dbs: 1 Gaps: 1

US-10-057-510-4 (1-216) x US-08-390-546-2 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTGTAAACACCTCACAGATTCCA 371
QY 21 AlasrGluGlnGluThrValArgProLysProLeuLeuLeuLeuLeuLeuLeuLeu 40
Db 372 GCTTCGGAACAAGAGACCTGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrValMetLysGluValLeuPheValThrLeuLeuLeuLeuLeu 59
Db 432 GTTGTGCAACAAAAGACACTTACTATGAAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuValProSerPheSerValLysGlnHisLeuValThrCysSerAsnAsp 79
Db 492 ATTATGACTAAACGATTATATGATGAGAAGCAACATATTGTTATTTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGlnHisArgLysIleValThr 99
Db 552 CTTCTAGGAGATTGTTGGCGTCCAGCTTCTGTGAAGAGAGACAGGAAATATAT 611
QY 100 ThrMetIleValArgLeuValValAsnGlnGlnLysSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGGAACCTTGTAGTCAATCAGCAGGAATCATCGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTAGTGAGAACAGGTTGACCTTTGAAGAGGTTGATCAAAAGGACCTTGTACAA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GACCTTCAGGAAGAGAAACCTTCTTCATCTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgAlaIleSerGluThrGluAsnSerAspGluLeuSerGlyGluArgGln 179

Db 792 AGAAGGAGCAATAGTAGACAGAGAAATTCAGATCAATATCTGGTGAACGCAA 851
Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAACGCCCAAAATCTGATGATTTCCCTTTCTTCATGAAGCCCTGGCTGTGT 911
Qy 200 ValIleArgGluIleCysGluArgSerSerSerSerGluSerThrGly 216
Db 912 GTATAAGGGAGATATGTTGTGAAGAGCAGTAGCAGTGAATCTACAGGG 962

RESULT 7

US-08-390-479A-2
; Sequence 2, Application US/08390479A
; Patent No. 5618921
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & WITCOFF, LTD.
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,479A
; FILING DATE: 02-FEB-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.48992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BEMB UT
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 2372 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL LINE: CaCo-2
; POSITION IN GENOME:
; MAP POSITION: 12q12-14
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 312..1784

US-08-390-479A-2

Alignment Scores:
Pred. No.: 2,126-115
Score: 1084.50
Percent Similarity: 99.54%
Best Local Similarity: 99.54%
Query Match: 99.04%
DB: 1
Length: 2372
Matches: 216
Conservative: 0
Mismatch: 0
Indels: 1
Gaps: 1

US-10-057-510-4 (1-216) x US-08-390-479A-2 (1-2372)
Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATCTGTGACCTACTGATGTGTGTGAACCACTCAGAGTTCCA 371
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLysProLeuLeuLeuLeuLysSer 40
Db 372 GCTTCGGAAACAGAGACCCCTGTTAGACCAAGCCATTGCTTTTGAAGTATTAAAGTCT 431
Qy 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu---GlnTy 59
Db 432 GTTGTGTCACAAAAGACACTTATATGAAAGAGGTCTTTTATCTTTGGCCAGTAT 491
Qy 60 IleMetThrLysArgLeuTyArgGluLysGlnGlnHisIleValTyCysSerAsnAsp 79
Db 492 ATTATGACTAACGATTATATGATGAGAGCAACACATATGTATATTGTTCAATGAT 551
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTy 99
Db 552 CTTCTAGGAGATTGTTTGGCGTGCCAAAGCTTCTCTGTGAAGAGCACAGAGAAATATAT 611
Qy 100 ThrMetIleTyArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGAACTTGTGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTGAGTGAGAACAGAGGTGTACCTTGAAGTGGGAGTGATCAAAAGAGCCCTTGACAA 731
Qy 140 GluLeuGlnGluLysProSerSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GAGCTTCAGAAAGAGAAACCTTCATCTCACATTGGTTTCTAGACCATCTACCTCATCT 791
Qy 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATATCTGTTGAACGACAA 851
Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAACGCCCAAAATCTGATGATTTCCCTTTCTTTGTATGAAGCCCTGCTGTGT 911
Qy 200 ValIleArgGluIleCysGluArgSerSerSerSerSerGluSerThrGly 216
Db 912 GTATAAGGGAGATATGTTGTGAAGAGCAGTAGCAGTGAATCTACAGGG 962

RESULT 8

US-08-557-393-2
; Sequence 2, Application US/08557393
; Patent No. 5762903
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,393


```

; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,500
; FILING DATE: 18-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42798
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2372 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL LINE: CaCo-2
; POSITION IN GENOME:
; MAP POSITION: 12q12-14
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 312..1784
; US-08-557-393-2

Alignment Scores:
Pred. No.: 2,12e-115 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 1 Gaps: 1

US-10-057-510-4 (1-216) x US-08-557-393-2 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACCTACTGTATGCTGTGTACCACTTACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysLeuLysSer 40
DB 372 GCTTCGGAACAGAGACCCTGGTTAGACCACCAAGCCATTGCTTTTGAAGTTATTAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrIleu---GlnTyr 59
DB 432 GTTGGTGCAAAAAGACACTTATCTATGTAAGAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 492 ATTATGACTTAACGATTATATGNTGGAAGCAACACATATTGTATATTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTTCTAGGAGATTGTTGGCGTCCCAAGCTTCTCTGTGAAAGAGACAGAGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAenGlnGlnLysSerAspSerGlyThr 119
DB 612 ACCATGATCTACAGGAACCTTGGTAGTCAATCAGCAGGAATCATCGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
DB 672 TCTGTGAGTGAGAACAGGTGTCCACTTGAAGGTGGGAGTGATCAAAAGGACCTTGTACAA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
DB 732 GAGCTTCAGGAAGAGAAACCTTCATCTTACATTTTGGTTTCTAGACCATCTACCTCATCT 791
```

Query Match: 99.04% Indels: 1
DB: 1 Gaps: 1
US-10-057-510-4 (1-216) x US-08-390-516C-2 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACTGATGATGCTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuSer 40
DB 372 GCTTCGGAAACAGACCTGTTAGACCAAGCCATGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 432 GTTGGTGACAAAAGACATTTACTATGAAAGAGGTTCTTTTATCTTGCCAGTAT 491
QY 60 IleMetThrLysArgLeuValAspGluLysGlnGlnHisLeValTyrCysSerAsnAsp 79
DB 492 ATTATGACTAAACGATTATATGATGACAAAGCAACACATATTTGATATTTGTTCAAAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTTCTAGGAGATTGTTTGGCGTCCAGCTTCTCTGTGAAGAGCACAGGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValValAsnGlnGlnGluSerSerGlyThr 119
DB 612 ACCATGATCTACAGAACTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
DB 672 TCTGTAGTGAGAACAGCTGTTAGACCAAGCCATGATCAAAAGGACCTTGTACAA 731
QY 140 GluLeuGlnGluGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
DB 732 GAGCTTTCAGGAGACAACTTCATCTTCATATTTGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyLysArgGln 179
DB 792 AGAAGGAGAGCAATAGTGAGACAGAAAGAAATTCAGATGAATATCTGCTGTAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerLysSerHisLeuValSerArgProSerThrSerSer 199
DB 852 AGAAACCCCAACAAATCTGATAGATATTTCCCTTTTCTTTGATGAAGCCCTGCTGTGT 911
QY 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
DB 912 GTAATAGGGAGATATGTTGTGAAGAGCAGTAGCAGTGAATCTACAGGG 962

RESULT 10

US-08-390-517A-2

Sequence 2, Application US/08390517A

Patent No. 5736338

GENERAL INFORMATION:

APPLICANT: BURRELL, MARILEE

APPLICANT: HILL, DAVID E.

APPLICANT: KINZLER, KENNETH W.

APPLICANT: VOGELSTEIN, BEET

TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDX2 GENE IN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

STREET: 1001 G STREET, N.W.

CITY: WASHINGTON

STATE: D.C.

COUNTRY: USA

ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,517A
FILING DATE: 07-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107,42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-3299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2372 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL LINE: CaCo-2
POSITION IN GENOME:
MAP POSITION: 12q12-14
FEATURE:
NAME/KEY: CDS
LOCATION: 312..1784
US-08-390-517A-2
Alignment Scores:
Pred. No.: 2,12e-115 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 1 Gaps: 1

US-10-057-510-4 (1-216) x US-08-390-517A-2 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACTGATGCTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuSer 40
DB 372 GCTTCGGAAACAGACCTGTTAGACCAAGCCATGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 432 GTTGGTGACAAAAGACATTTACTATGAAAGAGGTTCTTTTATCTTGCCAGTAT 491
QY 60 IleMetThrLysArgLeuValAspGluLysGlnGlnHisLeValTyrCysSerAsnAsp 79
DB 492 ATTATGACTAAACGATTATATGATGACAAAGCAACATATTTGATATTTGTTCAAAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTTCTAGGAGATTGTTTGGCGTCCAGCTTCTCTGTGAAGAGCACAGGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValValAsnGlnGlnGluSerSerGlyThr 119
DB 612 ACCATGATCTACAGAACTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
DB 672 TCTGTAGTGAGAACAGCTGTTAGACCAAGCCATGATCAAAAGGACCTTGTACAA 731
QY 140 GluLeuGlnGluGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
DB 732 GAGCTTTCAGGAGACAACTTCATCTTCATATTTGTTTCTAGACCATCTACCTCATCT 791

Qy 160 ArgArgAlaIleSerGluThrGluAenSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATTATCTGGTGAACGACAA 851
Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAAGCGCACAAATCTGATAGATATTTCCCTTTCTTTGATGAAGAGCCTGGCTCTGTGT 911
Qy 200 ValIleArgGluIleCysCysGluArgSerSerSerSerSerGluSerThrGly 216
Db 912 GTAATAGGGAGATATGTTGTGAAGAGCAGTAGCAGTGAATCTACAGGG 962

RESULT 11

US-08-390-515A-2
; Sequence 2, Application US/08390515A
; Patent No. 5756455
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,515A
; FILING DATE: 07-APR-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2372 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL LINE: CaCo-2
; POSITION IN GENOME:
; MAP POSITION: 12q12-14
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 312..1784
US-08-390-515A-2
Alignment Scores:
Pred. No.: 2,12e-115 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservatives: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1

Db: 1 Gaps: 1
US-10-057-510-4 (1-216) x US-08-390-515A-2 (1-2372)
Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATGTCGTGTACCTACTGATGGTGTGTAAACACCTCACGATTCCA 371
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLeuSer 40
Db 372 GCTTCGGACAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTATTAAAGTCT 431
Qy 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu---GlnTy 59
Db 432 GTTGTGTGCACAAAAGACACTTATACCTATGAAGAGGTCTCTTTTATCTTGGCCAGTAT 491
Qy 60 IleMetThrLysArgLeuTyArgGluLysGlnGlnHisIleValTyCysSerAsnAsp 79
Db 492 ATTATGACTAAACGATTATATGATGAGAGACACACATATTGTATATTGTTCAATGAT 551
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTy 99
Db 552 CTTCTAGGAGATTGTTTGGCGTCCCAAGCTTCTGTGAAAGAGACACAGAAAATATAT 611
Qy 100 ThrMetIleTyArgAsnLeuValValAsnGlnGlnGlnGlnGlnGlnGlnGlnGln 119
Db 612 ACCATGATCTACAGGAACCTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlyLysSerAspGlnLysAspLeuValGln 139
Db 672 TCTGTGAGTGAGACAGGTGTCCCTTGAAGTGGAGTGTATCAAAAGGACCTTGTACAA 731
Qy 140 GluLeuGlnGluLysProSerSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GAGCTTCAGGAAGAGAAACCTTCATCTTCACATTGCTTTCTAGACCATCTACCTCATCT 791
Qy 160 ArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATTAGTGAGACAGAAAGAAATTCAGATGAATATCTGTTGACGACAA 851
Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAAGCGCACAAATCTGATAGTATTTCCCTTTCTTTGATGAAGCCTGGCTCTGTGT 911

RESULT 12

US-08-801-718-2
; Sequence 2, Application US/08801718
; Patent No. 5858976
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

```

; APPLICATION NUMBER: US/08/801.718
; FILING DATE: 14-FEB-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/390,515
; FILING DATE: 07-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BEMB UT
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2372 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; CELL LINE: Caco-2
; POSITION IN GENOME:
; MAP POSITION: 12q12-14
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 312..1784
;
US-08-801-718-2
;
Alignment Scores:
Pred. No.: 2,12e-115 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 2 Gaps: 1
;
US-10-057-510-4 (1-216) x US-08-801-718-2 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSergInlePro 20
Db 312 ATGTGCAATACCAACATGCTGTACCTACTGCTGTACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysSer 40
Db 372 GCTTCGGAAACAAGAGACCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 432 GTTCGTGCACAAAAGACACTTACTATGAAAGAGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
Db 492 ATTATGACTAAACGATTATATGATGAGAACCAACACATATGTTATTTCTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 552 CTTCTAGAGATTTGTTTGGGTGCCAAGCTTCTCTGTGAAGAGACACAGGAAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAlaGlnGlnGlnGlnSerSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGAACTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTGAGTGAGACAGGTGTCACCTTGAAGGTGGAGTGATCAAAAGGACCTTGATCAA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
;
732 GAGCTTCAGGAAGAGAAACCTTCATCTTCACATTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGAGAGCAATTTAGTACAGACAGAGAAATTCAGATGAATTAATCTGGTGAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAAACGCCACAAATCTGATAGTATTTCCTTTCCTTTGATGAAGCGCTGCTCTGTGT 911
QY 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
Db 912 GTAATAAGGAGATATGTTGTGAAGAAGACGATAGCAGTGAATCTACAGGG 962
;
RESULT 13
US-09-073-567-1
; Sequence 1, Application US/09073567
; Patent No. 6013786
; GENERAL INFORMATION:
; APPLICANT: Jiaodong Chen
; APPLICANT: Sudhir Agrawal
; APPLICANT: Ruiwen Zhang
; TITLE OF INVENTION: MDM2-SPECIFIC ANTISENSE OLIGONUCLEOTIDES
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive, 32nd Floor
; CITY: Chicago
; STATE: IL
; COUNTRY: United States of America
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Microsoft Word 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,567
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Greenfield, Michael S.
; REGISTRATION NUMBER: 37,147
; REFERENCE/DOCKET NUMBER: 98,057-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 913-0001
; TELEFAX: (312) 913-0002
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2372 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: hmdm2 DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
US-09-073-567-1
;
Alignment Scores:
Pred. No.: 2,12e-115 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 3 Gaps: 1
;
US-10-057-510-4 (1-216) x US-09-073-567-1 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSergInlePro 20
Db 312 ATGTGCAATACCAACATGCTGTACCTACTGCTGTACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysSer 40
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:48:49 ; Search time 40.9655 Seconds
(without alignments)
1663.641 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 216

Sequence: 1 MONTNMSVPTDGAVTTSQIP.....ALCVIREICRRSSSESTG 216

Scoring table:

OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archesp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	130	60.2	135	Q96DS4	Q96ds4 homo sapien
2	120	55.6	446	Q8WYJ1	Q8wyj1 homo sapien
3	112	51.9	243	Q8TE47	Q8te47 homo sapien
4	89	41.2	166	Q8NDW2	Q8ndw2 homo sapien
5	57	26.4	66	Q96DS3	Q96ds3 homo sapien
6	57	26.4	69	Q86WA4	Q86wa4 homo sapien
7	57	26.4	70	Q86WA3	Q86wa3 homo sapien
8	57	26.4	105	Q8NDW0	Q8ndw0 homo sapien
9	57	26.4	118	Q8WYJ3	Q8wyj3 homo sapien
10	57	26.4	173	Q8TE46	Q8te46 homo sapien
11	57	26.4	436	Q8WYJ2	Q8wyj2 homo sapien
12	56	25.9	95	Q96DS1	Q96ds1 homo sapien
13	55	25.5	84	Q96DS2	Q96ds2 homo sapien
14	54	25.0	487	Q9GMZ6	Q9gmz6 canis faml
15	54	25.0	491	Q7YRZ8	Q7yrz8 felis silve
16	52	24.1	60	Q96DS5	Q96ds5 homo sapien

17	52	24.1	70	4	Q8NDW1	Q8ndw1 homo sapien
18	52	24.1	130	4	Q9H4C3	Q9h4c3 homo sapien
19	52	24.1	159	4	Q96DS0	Q96ds0 homo sapien
20	52	24.1	243	4	Q9H4C5	Q9h4c5 homo sapien
21	51	23.6	65	4	Q86WA5	Q86wa5 homo sapien
22	51	23.6	426	6	Q9GK41	Q9gk41 canis faml
23	48	22.2	50	4	Q8NDV9	Q8ndv9 homo sapien
24	40	18.5	54	4	Q86WA2	Q86wa2 homo sapien
25	29	13.4	489	11	Q91XK7	Q91xk7 m adult mal
26	26	12.0	98	4	Q9H4C2	Q9h4c2 homo sapien
27	20	9.3	151	4	Q9H4C4	Q9h4c4 homo sapien
28	11	5.1	325	13	Q9PVL2	Q9pvl2 gallus gall
29	9	4.2	1185	11	Q8CHA4	Q8cha4 mus musculu
30	8	3.7	151	16	Q8A5C7	Q8a5c7 bacteroides
31	8	3.7	185	5	Q9VPR1	Q9vpr1 drosophila
32	8	3.7	282	2	Q87877	Q87877 thauera aro
33	8	3.7	399	16	Q8A4U5	Q8a4u5 bacteroides
34	8	3.7	400	16	Q81EM6	Q81em6 bacillus ce
35	8	3.7	510	13	Q7SXZ6	Q7sxz6 brachydanio
36	8	3.7	522	10	Q23355	Q23355 arabidopsis
37	8	3.7	630	11	Q8B2J2	Q8b2j2 mus musculu
38	8	3.7	641	11	Q8BJG8	Q8bjg8 mus musculu
39	8	3.7	681	16	Q9L1X1	Q9l1x1 streptomyce
40	8	3.7	768	10	Q8L7I7	Q8l7i7 arabidopsis
41	8	3.7	876	10	Q9XF50	Q9xf50 arabidopsis
42	8	3.7	876	10	Q84K16	Q84k16 arabidopsis
43	8	3.7	910	10	Q9LRA3	Q9lra3 arabidopsis
44	8	3.7	1104	11	Q8BYH8	Q8byh8 mus musculu
45	8	3.7	1881	10	Q7XST8	Q7xst8 oryza sativ

ALIGNMENTS

RESULT 1

Q96DS4 PRELIMINARY; PRT; 195 AA.

AC Q96DS4; 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE MDM2 variant FB26.
 GN MDM2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN [1]
 RX NCBI_TaxID=9606;
 RC TISSUE=Rhabdomyosarcoma;
 RA Bartel F., Taylor A.C., Taubert H., Harris L.C.;
 RT "Novel mdm2 splice variants identified in pediatric rhabdomyosarcoma tumors and cell lines."
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF385323; AAL3243.1; -
 DR InterPro; IPR003121; SWIB.
 DR Pfam; PF02203; SWIB; 1
 SQ SEQUENCE 195 AA; 22161 MW; 4987AE567DB12DSD CRC64;

Query Match 60.2%; Score 130; DB 4; Length 195;

Best Local Similarity 100.0%; Pred. No. 5.2e-131;

Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	58	QYITKRLYDKQKHIVYCSNDLGLFGVPSFVKHRKIYTWIYNLVVNVQESDS	117
DB	59	QYITKRLYDKQKHIVYCSNDLGLFGVPSFVKHRKIYTWIYNLVVNVQESDS	118
QY	118	GTSVSENRCHEGSGDQDLVQELQEKPSSSHVSPSTSSRRRA;SETEENSELSGE	177
DB	119	GTSVSENRCHEGSGDQDLVQELQEKPSSSHVSPSTSSRRRA;SETEENSELSGE	178
QY	178	QRRKRHSKDS	187

```
Db 179 RQRKHKSDS 188
RESULT 2
Q8WYJ1 PRELIMINARY; PRT; 446 AA.
AC Q8WYJ1;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE MDM2 protein.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21248713; PubMed=11351297;
RA Tamborini E., Della Torre G., Lavarino C., Azzarelli A.,
RA Carpinelli P., Pierotti M.A., Pilotti S.;
RT "Analysis of the molecular species generated by MDM2 gene
RL Int. J. Cancer 92:790-796(2001).
DR EMBL; AF032845; XAL40180.1; -.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf_RanGDP.
DR Pfam; PFO2201; SWIB; 1.
DR Pfam; PFO641; zf-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS01359; ZF_RANBP2_2; 1.
DR PROSITE; PS50089; ZF_RING_1.
SQ SEQUENCE 446 AA; 49898 MW; 4B630B50750EADFE CRC64;
Query Match 55.6%; Score 120; DB 4; Length 446;
Best Local Similarity 100.0%; Pred. No. 6.2e-120; Indels 0; Gaps 0;
Matches 120; Conservative 0; Mismatches 0;
QY 97 KIYTMIRNLVNVNQESSDGTSTSVSENCHLEGGSDQKDLVQELQEKPSHLSVRPS 156
Db 53 KIYTMIRNLVNVNQESSDGTSTSVSENCHLEGGSDQKDLVQELQEKPSHLSVRPS 112
QY 157 TSSRRRAISETENSDELGSGRQRKHKSDSISLFDLSLALCVIREICCRSSSSSTG 216
Db 113 TSSRRRAISETENSDELGSGRQRKHKSDSISLFDLSLALCVIREICCRSSSSSTG 172
RESULT 3
Q8TE47 PRELIMINARY; PRT; 243 AA.
AC Q8TE47;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE MDM2 isoform K59 protein.
GN MDM2 ISOFORM K59.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Lymphocytes;
RA Bartel F., Pinkert D., Kappler M., Bache M., Schmidt H., Taubert H.;
RT "Alternatively and Aberrantly Spliced Transcripts of the MDM2-mRNA
RT Occur Frequently in Human Soft Tissue Sarcomas and in Multiple Normal
RT Tissues."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430612; CAD23251.1; -.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001841; Znf_Ring.
```

```
DR Pfam; PFO2201; SWIB; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 243 AA; 27317 MW; 9EB5D0142CF185A2 CRC64;
Query Match 51.9%; Score 112; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.4e-111;
Matches 112; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLVNVNQESSDS 117
Db 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLVNVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISSTEE 169
Db 119 GTSVSENCHLEGGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISSTEE 170
RESULT 4
Q8NDW2 PRELIMINARY; PRT; 166 AA.
AC Q8NDW2;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE P53-binding protein.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Bartel F., Pinkert D., Kappler M., Bache M., Schmidt H., Taubert H.;
RT "Alternatively and aberrantly spliced transcripts of the MDM2 mRNA
RT occur frequently in human soft tissue sarcoma and in multiple normal
RT tissues."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ491598; CAD36959.1; -.
DR InterPro; IPR003121; SWIB.
DR Pfam; PFO2201; SWIB; 1.
SQ SEQUENCE 166 AA; 18900 MW; FAG5BA18E85040D CRC64;
Query Match 41.2%; Score 89; DB 4; Length 166;
Best Local Similarity 100.0%; Pred. No. 5.1e-87;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLVNVNQESSDS 117
Db 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLVNVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQEK 146
Db 119 GTSVSENCHLEGGSDQKDLVQELQEK 147
RESULT 5
Q96DS3 PRELIMINARY; PRT; 66 AA.
AC Q96DS3;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE MDM2 variant FB28.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Rhabdomyosarcoma;
RA Bartel F., Taylor A.C., Taubert H., Harris L.C.;
```


RT "Novel mdm2 splice variants identified in pediatric rhabdomyosarcoma
tumors and cell lines."
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF365324; AAL13244.1; --
DR InterPro: IPR003121; SWIB.
DR Pfam: PF02201; SWIB; 1.
SQ SEQUENCE 66 AA; 7396 MW; E3B3F3C385D4A8A5 CRC64;

Query Match 26.4%; Score 57; DB 4; Length 66;
Best Local Similarity 100.0%; Pred. No. 5.3e-53;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL 57
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL 57

RESULT 6

Q86WA4 ID Q86WA4 PRELIMINARY; PRT; 69 AA.
AC Q86WA4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HDM2-HD3 protein.
GN HDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sturzenhocker B., Schlott T., Quentin T., Kube D., Jung W.,
RA Trumper L.;
RT "Abundant expression of spliced HDM2 in Hodgkin Lymphoma cells does
not interfere with p14ARF and p53 binding."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ550518; CAD79457.1; --
KW Alternative splicing.
SQ SEQUENCE 69 AA; 7625 MW; CAD81EC6B46E439A CRC64;

Query Match 26.4%; Score 57; DB 4; Length 69;
Best Local Similarity 100.0%; Pred. No. 5.5e-53;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL 57
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL 57

RESULT 7

Q86WA3 ID Q86WA3 PRELIMINARY; PRT; 70 AA.
AC Q86WA3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE HDM2-HD5 protein.
GN HDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sturzenhocker B., Schlott T., Quentin T., Kube D., Jung W.,
RA Trumper L.;
RT "Abundant expression of spliced HDM2 in Hodgkin Lymphoma cells does
not interfere with p14ARF and p53 binding."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ550519; CAD79458.1; --
KW Alternative splicing.
SQ SEQUENCE 70 AA; 7825 MW; 26ABFB61E3CF3629 CRC64;

Query Match 26.4%; Score 57; DB 4; Length 70;
Best Local Similarity 100.0%; Pred. No. 5.6e-53;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL 57
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL 57

RESULT 8

Q8NDW0 ID Q8NDW0 PRELIMINARY; PRT; 105 AA.
AC Q8NDW0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE P53-binding protein.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Bartel F., Pinkert D., Kappeler M., Bache M., Schmidt H., Taubert H.;
RT "Alternatively and aberrantly spliced transcripts of the MDM2 mRNA
occur frequently in human soft tissue sarcoma and in multiple normal
tissues."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ491700; CAD36961.1; --
DR InterPro: IPR003121; SWIB.
DR Pfam: PF02201; SWIB; 1.
SQ SEQUENCE 105 AA; 11940 MW; 289127D911672D63 CRC64;

Query Match 26.4%; Score 57; DB 4; Length 105;
Best Local Similarity 100.0%; Pred. No. 8.1e-53;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL 57
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL 57

RESULT 9

Q8WYJ3 ID Q8WYJ3 PRELIMINARY; PRT; 118 AA.
AC Q8WYJ3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MDM2 protein (fragment).
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21248713; PubMed=11351297;
RA Tamborini E., Della Torre G., Lavarino C., Azzarelli A.,
RA Carpinelli P., Pierotti M.A., Pilotti S.;
RT "Analysis of the molecular species generated by MDM2 gene
amplification in liposarcomas."
RL Int. J. Cancer 92:790-796(2001).
DR EMBL: AF092843; AAL40178.1; --
DR InterPro: IPR003121; SWIB.
DR Pfam: PF02201; SWIB; 1.
FT NON_TER 118 118
SQ SEQUENCE 118 AA; 13536 MW; D7A4DBAA93D8841B CRC64;

Query Match 26.4%; Score 57; DB 4; Length 118;

Best Local Similarity 100.0%; Pred. No. 9e-53;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLFYL 57
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLFYL 57

RESULT 10
Q8TE46 PRELIMINARY; PRT; 173 AA.
AC Q8TE46;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
GN MDM2 isoform N1.40 protein.
GN MDM2 ISOFORM N1.40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Bartel F.; Pinkert D.; Kappler M.; Bache M.; Schmidt H.; Taubert H.;
RT "Alternatively and Aberrantly Spliced Transcripts of the MDM2-mRNA
RT Occur Frequently in Human Soft Tissue Sarcomas and in Multiple Normal
RT Tissues.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ430614; CAD23252.1; -
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001841; Znf_fing.
DR Pfam; PF02201; SWIB; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 173 AA; 19508 MW; AB26BECFA003B261 CRC64;

Query Match 26.4%; Score 57; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 1.3e-52;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLFYL 57
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLFYL 57

RESULT 11
Q8WYJ2 PRELIMINARY; PRT; 436 AA.
AC Q8WYJ2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE MDM2 protein.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21248713; PubMed=11351297;
RA Tamborini E.; Della Torre G.; Lavarino C.; Azzarelli A.;
RA Carpinelli P.; Pierotti M.A.; Pilotti S.;
RT "Analysis of the molecular species generated by MDM2 gene
RT amplification in liposarcomas.";
RL Int. J. Cancer 92:790-796(2001).
DR EMBL; AF092844; AAL40179.1; -
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf_RangRP.
DR InterPro; IPR001841; Znf_fing.
DR Pfam; PF02201; SWIB; 1.
DR Pfam; PF00641; zf-RanBP; 1.

SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS01359; ZF_RANBP2_2; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 436 AA; 49248 MW; 3C9F55E98BC4203A CRC64;

Query Match 26.4%; Score 57; DB 4; Length 436;
Best Local Similarity 100.0%; Pred. No. 3e-52;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLFYL 57
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLFYL 57

RESULT 12
Q96DS1 PRELIMINARY; PRT; 95 AA.
AC Q96DS1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE MDM2 variant PB30.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Rhabdomyosarcoma;
RA Bartel F.; Taylor A.C.; Taubert H.; Harris L.C.;
RT "Novel mdm2 splice variants identified in pediatric rhabdomyosarcoma
RT tumors and cell lines.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF385326; AAL13246.1; -
DR InterPro; IPR003121; SWIB.
DR Pfam; PF02201; SWIB; 1.
SQ SEQUENCE 95 AA; 10622 MW; 00052F95211E3612 CRC64;

Query Match 25.9%; Score 56; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 8.7e-52;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLFYL 56
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLFYL 56

RESULT 13
Q96DS2 PRELIMINARY; PRT; 84 AA.
AC Q96DS2;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE MDM2 variant PB29.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Rhabdomyosarcoma;
RA Bartel F.; Taylor A.C.; Taubert H.; Harris L.C.;
RT "Novel mdm2 splice variants identified in pediatric rhabdomyosarcoma
RT tumors and cell lines.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF385325; AAL13245.1; -
SQ SEQUENCE 84 AA; 8955 MW; 6BA9AE1E684F1D31 CRC64;

Query Match 25.5%; Score 55; DB 4; Length 84;

Best Local Similarity 100.0%; Pred. No. 9,3e-51;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 113 ESSDSGTSVSENCHLEGSDQKDLVQELQEEKPSSSHLYSRPSTSSRRRAISET 167
Db 28 ESSDSGTSVSENCHLEGSDQKDLVQELQEEKPSSSHLYSRPSTSSRRRAISET 82

RESULT 14

Q9GMZ6 PRELIMINARY; PRT; 487 AA.
AC Q9GMZ6; 2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE MDM2.
GN MDM2.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Setoguchi A., Tsujimoto H.;
RT "Dog mdm2 cDNA."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB031276; BAB11975.1; -
DR HSP; Q9UNT8; 1YCR.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf_RanGDP.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF02201; SWIB; 1.
DR Pfam; PF00641; zf-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS50193; ZF_RANBP2_2; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 487 AA; 54724 MW; 34FC5CC6A18D7744 CRC64;

Query Match 25.0%; Score 54; DB 6; Length 487;
Best Local Similarity 100.0%; Pred. No. 5.5e-49;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 QYIMTKRLYDEKQOHVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVWVWQ 111
Db 59 QYIMTKRLYDEKQOHVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVWVWQ 112

RESULT 15

Q7YRZ8 PRELIMINARY; PRT; 491 AA.
AC Q7YRZ8;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Double minute 2 protein MDM2.
GN MDM2.
OS Felis silvestris catus (Cat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
OX NCBI_TaxID=9685;
RN [1]
RP SEQUENCE FROM N.A.
RA Mixi R., Okuda M., Ma Z., Inokuma H., Onishi T.;
RT "Molecular cloning of feline mdm2 cDNA."
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB099709; BAC78209.1; -
SQ SEQUENCE 491 AA; 55433 MW; D93E25D5638E8934 CRC64;

Query Match 25.0%; Score 54; DB 6; Length 491;
Best Local Similarity 100.0%; Pred. No. 5.6e-49;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 QYIMTKRLYDEKQOHVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVWVWQ 111
Db 59 QYIMTKRLYDEKQOHVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVWVWQ 112

Search completed: March 14, 2004, 21:54:47
Job time : 41.9655 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:42:43 ; Search time 28.3034 Seconds
(without alignments)
393.988 Million cell updates/sec

Title: US-10-057-510-4
Perfect score: 1095
Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIREICCRSSSSSESTG 216

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/2/iaa/6C_COMB.pep.*
- 6: /cgn2_6/prodata/2/iaa/6D_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1095	100.0	216	US-09-510-252-4	Sequence 4, Appli
2	1084.5	99.0	491	US-07-903-103-2	Sequence 2, Appli
3	1084.5	99.0	491	US-08-044-619A-2	Sequence 2, Appli
4	1084.5	99.0	491	US-08-283-911-2	Sequence 2, Appli
5	1084.5	99.0	491	US-08-245-500A-3	Sequence 3, Appli
6	1084.5	99.0	491	US-08-390-546-3	Sequence 3, Appli
7	1084.5	99.0	491	US-08-390-479A-3	Sequence 3, Appli
8	1084.5	99.0	491	US-08-557-393-3	Sequence 3, Appli
9	1084.5	99.0	491	US-08-390-516C-3	Sequence 3, Appli
10	1084.5	99.0	491	US-08-390-517A-3	Sequence 3, Appli
11	1084.5	99.0	491	US-08-390-515A-3	Sequence 3, Appli
12	1084.5	99.0	491	US-08-801-718-3	Sequence 3, Appli
13	1084.5	99.0	491	US-09-170-159A-3	Sequence 3, Appli
14	1084.5	99.0	491	US-09-480-718-4	Sequence 4, Appli
15	940.5	85.9	188	US-09-603-052-4	Sequence 4, Appli
16	804.5	73.5	489	US-07-903-103-4	Sequence 4, Appli
17	804.5	73.5	489	US-08-044-619A-4	Sequence 4, Appli
18	804.5	73.5	489	US-08-283-911-4	Sequence 4, Appli
19	804.5	73.5	489	US-08-245-500A-5	Sequence 5, Appli
20	804.5	73.5	489	US-08-390-546-5	Sequence 5, Appli
21	804.5	73.5	489	US-08-390-479A-5	Sequence 5, Appli
22	804.5	73.5	489	US-08-557-393-5	Sequence 5, Appli
23	804.5	73.5	489	US-08-390-516C-5	Sequence 5, Appli
24	804.5	73.5	489	US-08-390-517A-5	Sequence 5, Appli
25	804.5	73.5	489	US-08-390-515A-5	Sequence 5, Appli
26	804.5	73.5	489	US-08-801-718-5	Sequence 5, Appli
27	804.5	73.5	489	US-09-170-159A-5	Sequence 5, Appli

Sequence 46, Appli
Sequence 6, Appli
Sequence 3, Appli
Sequence 7, Appli
Sequence 46, Appli
Sequence 46, Appli
Sequence 28, Appli
Sequence 5, Appli
Sequence 5, Appli
Sequence 78, Appli
Sequence 78, Appli
Sequence 78, Appli
Sequence 78, Appli
Sequence 3, Appli
Sequence 15, Appli
Sequence 15, Appli
Sequence 4, Appli

ALIGNMENTS

RESULT 1

US-09-510-252-4
; Sequence 4, Application US/09510252
; Patent No. 6372490
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Mei-jia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDM INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT APPLICATION NUMBER: US/09/510,252
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-510-252-4

Query Match 100.0%; Score 1095; DB 4; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.1e-107;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MCNTNMSVPTDGAVTTSQIPASQEITLVRPKPLLLKLLKSVGAOKDTYTWKEVLFYQYI	60
Db	1	MCNTNMSVPTDGAVTTSQIPASQEITLVRPKPLLLKLLKSVGAOKDTYTWKEVLFYQYI	60
QY	61	MTKELYDEKQOHVYVCSNDLLGLDFGVPSFVKHRIKTYMYRNLVVNNQSSDSGTS	120
Db	61	MTKELYDEKQOHVYVCSNDLLGLDFGVPSFVKHRIKTYMYRNLVVNNQSSDSGTS	120
QY	121	VSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSPSSRRRAJSETEENSDELSEGERQ	180
Db	121	VSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSPSSRRRAJSETEENSDELSEGERQ	180
QY	181	KRKHSDSISLFSDESALCVIREICCRSSSSSESTG	216
Db	181	KRKHSDSISLFSDESALCVIREICCRSSSSSESTG	216

RESULT 2

US-07-903-103-2
; Sequence 2, Application US/07903103
; Patent No. 5411860

incorrect

GENERAL INFORMATION:
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
TITLE OF INVENTION: HUMAN TUMORS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G ST., N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4597

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903.103
FILING DATE: 07-APR-1992
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 32,141
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 01107.40148
REFERENCE/DOCKET NUMBER: 01107.40148

TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-07-903-103-2
Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 60

QY 60 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 120

QY 120 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 179
DB 121 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 180

QY 180 RKRHKSISLSPDESALCVIREICCRSSSESTG 216
DB 181 RKRHKSISLSPDESALCVIREICCRSSSESTG 217

RESULT 3
US-08-044-619A-2
Sequence 2, Application US/08044619A
Patent No. 542063

GENERAL INFORMATION:
APPLICANT: THE JOHNS HOPKINS UNIVERSITY
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
TITLE OF INVENTION: HUMAN TUMORS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G ST., N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA

ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/903.103
FILING DATE: 07-APR-1992
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 32,141
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 01107.40148
REFERENCE/DOCKET NUMBER: 01107.40148

TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-07-903-103-2
Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 60

QY 60 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 120

QY 120 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 179
DB 121 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 180

QY 180 RKRHKSISLSPDESALCVIREICCRSSSESTG 216
DB 181 RKRHKSISLSPDESALCVIREICCRSSSESTG 217

RESULT 3
US-08-044-619A-2
Sequence 2, Application US/08044619A
Patent No. 542063

GENERAL INFORMATION:
APPLICANT: THE JOHNS HOPKINS UNIVERSITY
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
TITLE OF INVENTION: HUMAN TUMORS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G ST., N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA

ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/044.619A
FILING DATE: 07-APR-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903.103
FILING DATE: 23-JUN-1992
APPLICATION NUMBER: US 07/867.840
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.40148
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-044-619A-2
Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G ST., N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20001-4597

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/044.619A
FILING DATE: 07-APR-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903.103
FILING DATE: 23-JUN-1992
APPLICATION NUMBER: US 07/867.840
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.40148
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-044-619A-2
Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 60

QY 60 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 120

QY 120 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 179
DB 121 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 180

QY 180 RKRHKSISLSPDESALCVIREICCRSSSESTG 216
DB 181 RKRHKSISLSPDESALCVIREICCRSSSESTG 217

RESULT 4
US-08-283-911-2
Sequence 2, Application US/08283911
Patent No. 5519118

GENERAL INFORMATION:
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
TITLE OF INVENTION: HUMAN TUMORS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G ST., N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA

ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/044.619A
FILING DATE: 07-APR-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903.103
FILING DATE: 23-JUN-1992
APPLICATION NUMBER: US 07/867.840
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.40148
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-044-619A-2
Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 60

QY 60 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 120

QY 120 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 179
DB 121 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 180

QY 180 RKRHKSISLSPDESALCVIREICCRSSSESTG 216
DB 181 RKRHKSISLSPDESALCVIREICCRSSSESTG 217

RESULT 4
US-08-283-911-2
Sequence 2, Application US/08283911
Patent No. 5519118

GENERAL INFORMATION:
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
TITLE OF INVENTION: HUMAN TUMORS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:

ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G ST., N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA

ZIP: 20001-4597
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/044.619A
FILING DATE: 07-APR-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/903.103
FILING DATE: 23-JUN-1992
APPLICATION NUMBER: US 07/867.840
FILING DATE: 07-APR-1992
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.40148
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-044-619A-2
Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRKPLLLKLLKSVGAQKDTYTKVLFYLY-QY 60

QY 60 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQKHIVYCSNDLLGLFGVPSPSVKHKIYTWIYRNLVNVNQESSDSGT 120

QY 120 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 179
DB 121 SVSENRCHEGGSQKDLVQELQBEKPSHLSVSRPSTSSRRRAISETEENSDELGERQ 180

QY 180 RKRHKSISLSPDESALCVIREICCRSSSESTG 216
DB 181 RKRHKSISLSPDESALCVIREICCRSSSESTG 217

```

; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,500A
; FILING DATE: 07-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-245-500A-3

Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDTYTMKEVLFYLGQY 60
Qy 60 IMTKRLYDEKQHHVYCSNDLLGDLFGVPSFVKHEHKIYTMIRNVLVNVNQESSDSTG 119
Db 61 IMTKRLYDEKQHHVYCSNDLLGDLFGVPSFVKHEHKIYTMIRNVLVNVNQESSDSTG 120
Qy 120 SVSENCHLEGGSDQKDLVQELQEKPSSSHLSVSRPSTSSRRRAISETEENSDELGSRQ 179
Db 121 SVSENCHLEGGSDQKDLVQELQEKPSSSHLSVSRPSTSSRRRAISETEENSDELGSRQ 180
Qy 180 RKHKSDSLSLSPDESALCVIREICCCRSSSESTG 216
Db 181 RKHKSDSLSLSPDESALCVIREICCCRSSSESTG 217

RESULT 5
; Sequence 3, Application US/08245500A-3
; Patent No. 5550023
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,546
; FILING DATE: 07-APR-1993
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:

```

NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-546-3

Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDTYTMKEVLFYLGQY 59
Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDTYTMKEVLFYLGQY 60
QY 60 IMTKELYDEKQKHIVYCSNDLLGLFGVPSPSVKHEHRIYTMIRNLVNVNQESSDST 119
Db 61 IMTKELYDEKQKHIVYCSNDLLGLFGVPSPSVKHEHRIYTMIRNLVNVNQESSDST 120
QY 120 SVSENRCHEGGSDQKDLVQELQBEKPSHLSVRPSTSSRRRAISETEENSDELGERQ 179
Db 121 SVSENRCHEGGSDQKDLVQELQBEKPSHLSVRPSTSSRRRAISETEENSDELGERQ 180
QY 180 RKRHKSISLSFDESALCVIREICCRSSSSESTG 216
Db 181 RKRHKSISLSFDESALCVIREICCRSSSSESTG 217

RESULT 7

US-08-390-479A-3
Sequence 3, Application US/08390479A
Patent No. 5618921
GENERAL INFORMATION:
APPLICANT: BURRELL, MARILEE
APPLICANT: HILL, DAVID E.
APPLICANT: KINZLER, KENNETH W.
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER & WITCOFF, LTD.
STREET: 1001 G STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,479A
FILING DATE: 02-FEB-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.48992
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-479A-3

Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDTYTMKEVLFYLGQY 59
Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDTYTMKEVLFYLGQY 60
QY 60 IMTKELYDEKQKHIVYCSNDLLGLFGVPSPSVKHEHRIYTMIRNLVNVNQESSDST 119
Db 61 IMTKELYDEKQKHIVYCSNDLLGLFGVPSPSVKHEHRIYTMIRNLVNVNQESSDST 120
QY 120 SVSENRCHEGGSDQKDLVQELQBEKPSHLSVRPSTSSRRRAISETEENSDELGERQ 179
Db 121 SVSENRCHEGGSDQKDLVQELQBEKPSHLSVRPSTSSRRRAISETEENSDELGERQ 180
QY 180 RKRHKSISLSFDESALCVIREICCRSSSSESTG 216
Db 181 RKRHKSISLSFDESALCVIREICCRSSSSESTG 217

RESULT 8

US-08-557-393-3
Sequence 3, Application US/08557393
Patent No. 5702903
GENERAL INFORMATION:
APPLICANT: BURRELL, MARILEE
APPLICANT: HILL, DAVID E.
APPLICANT: KINZLER, KENNETH W.
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,393
FILING DATE: 13-NOV-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/245,500
FILING DATE: 18-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-557-393-3

Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 60
QY 60 IMTKRLYDEKQOHIVYCSNDLLGLDGFVPSFVSKEHRKIYTMIRNVLVNVNQESSDSTG 119
DB 61 IMTKRLYDEKQOHIVYCSNDLLGLDGFVPSFVSKEHRKIYTMIRNVLVNVNQESSDSTG 120
QY 120 SVSNRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSEENDELSGRQ 179
DB 121 SVSNRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSEENDELSGRQ 180
QY 180 RKRHKSISLSPFDESALCVIREICERSSSSSESTG 216
DB 181 RKRHKSISLSPFDESALCVIREICERSSSSSESTG 217

RESULT 9

US-08-390-516C-3
; Sequence 3, Application US/08390516C
; Patent No. 5708136
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,516C
FILING DATE: 07-APR-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-516C-3

Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 60
QY 60 IMTKRLYDEKQOHIVYCSNDLLGLDGFVPSFVSKEHRKIYTMIRNVLVNVNQESSDSTG 119
DB 61 IMTKRLYDEKQOHIVYCSNDLLGLDGFVPSFVSKEHRKIYTMIRNVLVNVNQESSDSTG 120
QY 120 SVSNRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSEENDELSGRQ 179
DB 121 SVSNRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSEENDELSGRQ 180
QY 180 RKRHKSISLSPFDESALCVIREICERSSSSSESTG 216
DB 181 RKRHKSISLSPFDESALCVIREICERSSSSSESTG 217

DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 60
QY 60 IMTKRLYDEKQOHIVYCSNDLLGLDGFVPSFVSKEHRKIYTMIRNVLVNVNQESSDSTG 119
DB 61 IMTKRLYDEKQOHIVYCSNDLLGLDGFVPSFVSKEHRKIYTMIRNVLVNVNQESSDSTG 120
QY 120 SVSNRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSEENDELSGRQ 179
DB 121 SVSNRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSEENDELSGRQ 180
QY 180 RKRHKSISLSPFDESALCVIREICERSSSSSESTG 216
DB 181 RKRHKSISLSPFDESALCVIREICERSSSSSESTG 217

RESULT 10

US-08-390-517A-3
; Sequence 3, Application US/08390517A
; Patent No. 5736338
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; TITLE OF INVENTION: HUMAN TUMORS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,517A
FILING DATE: 07-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BBMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-517A-3

Query Match 99.0%; Score 1084.5; DB 1; Length 491;
Best Local Similarity 99.5%; Pred. No. 1.7e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 60
QY 60 IMTKRLYDEKQOHIVYCSNDLLGLDGFVPSFVSKEHRKIYTMIRNVLVNVNQESSDSTG 119
DB 61 IMTKRLYDEKQOHIVYCSNDLLGLDGFVPSFVSKEHRKIYTMIRNVLVNVNQESSDSTG 120
QY 120 SVSNRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSEENDELSGRQ 179
DB 121 SVSNRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSEENDELSGRQ 180
QY 180 RKRHKSISLSPFDESALCVIREICERSSSSSESTG 216
DB 181 RKRHKSISLSPFDESALCVIREICERSSSSSESTG 217

121	SVENRCHLEGGSDOKDVLQVLEKEKSSHLVSRPSTSSRRRAISZETSENSDELUGERQ	180
Db		
160	KRKHSDISLSFDESLALCVIRICERRSSSESTG	216
QY		
181	KRKHSDISLSFDESLALCVIRICERRSSSESTG	217
Db		

RESULT 11
 US-08-390-515A-3
 ; Sequence 3, Application US/08390515A
 ; Patent No. 5756455
 ; GENERAL INFORMATION:
 ; APPLICANT: BURRELL, MARILEE
 ; APPLICANT: HILL, DAVID E.
 ; APPLICANT: KINZLER, KENNETH W.
 ; APPLICANT: VOGELSTEIN, BERT
 ; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
 ; TITLE OF INVENTION: HUMAN TUMORS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
 ; STREET: 1001 G STREET, N.W.
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA

Query Match	99.0%;	Score 1084.5;	DB 1;	Length 491;
Best Local Similarity	99.5%;	Pred. No. 1.7e-105;		
Matches 216;	Conservative 0;	Mismatches 0;	Indels 1;	Gaps 1;
Qy	1	MCNTNMSVPTDGAVTTSQIQIPASBQETLVRPKPILLKLLKLSVGAKQDITYTMKEVFLV-L-QY	59	
Db	1	MCNTNMSVPTDGAVTTSQIQIPASBQETLVRPKPILLKLLKLSVGAKQDITYTMKEVFLVQLQY	60	
Qy	60	INTKRLYDEKQKHIVYCSNDLLDGLDFGVPFSVKRHKIYTMIRNLVNVNQESDSDGT	119	
Db	61	INTKRLYDEKQKHIVYCSNDLLDGLDFGVPFSVKRHKIYTMIRNLVNVNQESDSDGT	120	
Qy	120	SVSNRCHLEGGSDQKDLVQLQEKPSSSHLSVRPSTSSRRRAISETENDELSGSRQ	179	
Db	121	SVSNRCHLEGGSDQKDLVQLQEKPSSSHLSVRPSTSSRRRAISETENDELSGSRQ	180	
Qy	180	RKRKXSDISISFDESIALCVIREICCRSSSSSESTG	216	
Db	181	RKRKXSDISISFDESIALCVIREICCRSSSSSESTG	217	

US-08-801-718-3
Sequence 3, Application US/08801718
Patent No. 5858976
GENERAL INFORMATION:
APPLICANT: BURRELL, MARILEE
APPLICANT: HILL, DAVID E.
APPLICANT: KINZLER, KENNETH W.
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
TITLE OF INVENTION: HUMAN TUMORS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,718
FILING DATE: 14-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/390,515
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-718-3

Query Match	99.0%;	Score 1084.5;	DB 2;	Length 491;																											
Best Local Similarity	99.5%;	Pred. No. 1.7e-105;																													
Matches 216;	Conservative 0;	Mismatches 10;	Indels 1;	Gaps 1;																											
Qy	1	MONTNMSVPTDGA	VTT	SOIPASBQETLVR	PKP	LLKLLKLSVGAQKDTY	TMKEVFLVY	-QY	59																						
Db	1	MONTNMSVPTDGA	VTT	SOIPASBQETLVR	PKP	LLKLLKLSVGAQKDTY	TMKEVFLVY	LQY	60																						
Qy	60	IMTKRLYD	EKQ	HIVY	CSNDLL	LDGLFGVPS	F	SVKEHRKIY	TM	YRNLVYNQ	ESSD	S	D	S	G	T	119														
Db	61	IMTKRLYD	EKQ	HIVY	CSNDLL	LDGLFGVPS	F	SVKEHRKIY	TM	YRNLVYNQ	ESSD	S	D	S	G	T	120														
Qy	120	SVSENNCH	LEGG	S	D	KDLVQ	ELQ	EKP	SPSS	SHLY	SRP	ST	SS	RRRAI	SE	T	E	N	S	D	E	L	S	G	E	R	Q	179			
Db	121	SVSENNCH	LEGG	S	D	KDLVQ	ELQ	EKP	SPSS	SHLY	SRP	ST	SS	RRRAI	SE	T	E	N	S	D	E	L	S	G	E	R	Q	180			
Qy	180	RKEHKSD	S	I	S	L	S	P	D	E	S	L	A	L	C	V	R	E	I	C	E	R	S	S	S	S	E	S	T	G	216
Db	181	RKEHKSD	S	I	S	P	D	E	S	L	A	L	C	V	R	E	I	C	E	R	S	S	S	S	E	S	T	G	217		

RESULT 13
US-09-170-159A-3
; Sequence 3, Application US/09170159A
; Patent No. 6399755
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE

OF INVENTION: ART-23, A NOVEL REGISTRATION OF THE PATENT
REFERENCE: 1340-1-023 CIP 1
TIT APPLICATION NUMBER: US/09/480,718

QY 180 RRRHKS DS 187

Db |||||||
 181 RRRKSDS 188

Search completed: March 14, 2004, 21:49:42
Job time : 29.3034 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:49:49 ; Search time 21.6 Seconds
(without alignments)
961.915 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 216

Sequence: 1 MONTNMSVPTDGAVTTSQIP.....ALCVIREICCRSSSSSTG 216

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 78:*

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	159	73.6	491	1 S24354	p53-binding protei
2	29	13.4	489	2 S15349	mdm2 protein - mou
3	8	3.7	162	2 F71413	probable S-linaloo
4	8	3.7	363	2 D49889	Hrpl 3'-region hyp
5	8	3.7	522	2 B85165	hypothetical prote
6	7	3.2	100	2 AF1490	hypothetical prote
7	7	3.2	104	2 AC1130	hypothetical prote
8	7	3.2	105	2 H97030	probable membrane
9	7	3.2	106	2 T44366	hypothetical prote
10	7	3.2	133	2 S14234	hypothetical prote
11	7	3.2	133	2 H72476	hypothetical prote
12	7	3.2	149	2 JCS097	ferric uptake regu
13	7	3.2	151	2 T10768	latex allergen hev
14	7	3.2	153	2 E82330	conserved hypothet
15	7	3.2	187	2 T27416	hypothetical prote
16	7	3.2	196	2 A28332	alpha-crystallin c
17	7	3.2	235	2 T41107	hypothetical serin
18	7	3.2	237	2 AB1668	phosphoribosylamin
19	7	3.2	237	2 AD1296	phosphoribosylamin
20	7	3.2	266	2 D75545	phosphoribosylform
21	7	3.2	271	2 E71821	dimethyladenosine
22	7	3.2	278	2 C83552	hypothetical prote
23	7	3.2	296	2 F84092	hypothetical prote
24	7	3.2	345	2 A48462	dense granule prot
25	7	3.2	362	1 ZPBQC1	ubiquinol-cytochro
26	7	3.2	376	2 C362	hypothetical prote
27	7	3.2	380	2 T00683	hypothetical prote
28	7	3.2	400	2 T35334	probable membrane
29	7	3.2	408	2 C70685	hypothetical prote

30 7 3.2 416 1 ACCH3N
31 7 3.2 439 2 T16414
32 7 3.2 440 2 E84118
33 7 3.2 458 2 A43549
34 7 3.2 477 2 B95130
35 7 3.2 477 2 H98000
36 7 3.2 495 2 S60589
37 7 3.2 499 2 A24572
38 7 3.2 502 2 A37040
39 7 3.2 503 2 A53956
40 7 3.2 511 2 S58399
41 7 3.2 527 2 A83453
42 7 3.2 581 2 B83790
43 7 3.2 585 2 S48950
44 7 3.2 596 1 A44072
45 7 3.2 646 2 S56201

ALIGNMENTS

RESULT 1

S24354

p53-binding protein mdm2 - human

N:Alternate names: mdm-2 oncogene; mouse double minute 2 homolog; p53-associated phosph

C:Species: Homo sapiens (man)

C>Date: 17-Mar-2000 #sequence revision 17-Mar-2000 #text_change 17-Mar-2000

C:Accession: S24354; S57338; G02026

R:Oliner, J. D.; Kinzler, K. W.; Meltzer, P. S.; George, D. L.; Vogelstein, B.

Nature 358, 80-83, 1992

Article: Amplification of a gene encoding a p53-associated protein in human sarcomas.

A:Reference number: S24354; UID:92310576; PMID:1614537

A:Accession: S24354

A:Molecule type: mRNA

A:Residues: 1-491 <OI>

A:Cross-references: EMBL:Z12020; NID:G35211; PIDN:CAA78055.1; PID:G35212

R:Zauberman, A.; Flusberg, D.; Haupt, Y.; Barak, Y.; Oren, M.

Nucleic Acids Res. 23, 2584-2592, 1995

A:Title: A functional p53-responsive intronic promoter is contained within the human md

A:Reference number: S57338; UID:95380270; PMID:7651818

A:Accession: S57338

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-16, 'P', 18-24 <ZAU>

A:Cross-references: EMBL:U28935; NID:G904033; PIDN:AAA82237.1; PID:G904034

R:Lunec, J.

A:Description: Multiple alternate spliced mdm2 transcripts with loss of p53 binding dom

A:Reference number: G09070

A:Accession: G02026

A>Status: translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-27,223-491 <LUN>

A:Cross-references: EMBL:U33199; NID:G992676; PIDN:AAA75514.1; PID:G992677

A:Experimental source: splice form A

C:Genetics:

A:Gene: GDB:MDM2

A:Cross-references: GDB:250456; OMIM:164785

A:Map position: 12q14.3-12q15

C:Superfamily: human p53-binding protein mdm2

C:Keywords: alternative splicing; oncogene; phosphoprotein

F:1-491/Product: p53-binding protein mdm-2 #status predicted <MAT1>

F:1-27,223-491/Product: p53-binding protein mdm-2, splice form A #status predicted <MAT>

Query Match 73.6%; Score 159; DB 1; Length 491;

Best local similarity 100.0%; Pred. No. 5.5e-155;

Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFSVKEHKIYTWIYENLVVNNQSSSDS 117

Db 59 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFSVKEHKIYTWIYENLVVNNQSSSDS 118

QY 118 GTSVSENCHLEGGSDKDLVQLQELQEKFPSSHLVSPSTSSRRRAISETEENSDELSGE 177
DB 119 GTSVSENCHLEGGSDKDLVQLQELQEKFPSSHLVSPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKRHKSDSLSLFDLSLALCVIRECCERSSSSESTG 216
DB 179 RQRKRHKSDSLSLFDLSLALCVIRECCERSSSSESTG 217

RESULT 2
S15349
mdm2 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Mar-2000
C:Accession: S15349
R: Fakhrazadeh, S.S.; Trusko, S.P.; George, D.L.
EMBO J. 10, 1565-1569, 1991
A:Title: Tumorigenic potential associated with enhanced expression of a gene that is amp
A:Reference number: S15349; MUID:91224107; PMID:2026149
A:Accession: S15349
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-489 <FAK>
A:CROSS-references: ENBL:X58876; NID:G53038; PIDN:CAA41684.1; PID:G53039
C:Genetics:
A:Gene: mdm2
C:Superfamily: human p53-binding protein mdm2

Query Match 13.4%; Score 29; DB 2; Length 489;
Best Local Similarity 100.0%; Pred. No. 2e-21;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 TSQIPASEQETLVKPKLLKLSVGAQ 44
DB 16 TSQIPASEQETLVKPKLLKLSVGAQ 44

RESULT 3
F71413
Probable S-linalool synthase - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C:Accession: F71413
R: Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; And
C.; Chaiwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: F71413
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-162 <BEV>
A:CROSS-references: GB:Z57337; NID:G2244829; PID:G2244858
C:Genetics:
A:Map position: 4COP9-4G3845
C:Superfamily: Arabidopsis probable S-linalool synthase

Query Match 3.7%; Score 8; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 RQRKRHKHS 185
DB 69 RQRKRHKHS 76

RESULT 4
D49889

Hrpl 3'-region hypothetical protein - Pseudomonas syringae pv. syringae (fragment)
C:Species: Pseudomonas syringae pv. syringae
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 11-Aug-1995
C:Accession: D49889
R: Xiao, Y.; Heu, S.; Yi, J.; Lu, Y.; Hutcheson, S.W.
J. Bacteriol. 176, 1025-1036, 1994
A:Title: Identification of a putative alternate sigma factor and characterization of a
hnaA genes.
A:Reference number: A49889; MUID:94148760; PMID:8106313
A:Accession: D49889
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-363 <XIA>
A:CROSS-references: GB:U03855
C:Genetics:
A:Gene: hrpK

Query Match 3.7%; Score 8; DB 2; Length 363;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 RSSSEST 215
DB 74 RSSSEST 81

RESULT 5
B85165
hypothetical protein dl3551w [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: B85165
R: anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spri
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: B85165
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-522 <STO>
A:CROSS-references: GB:NC_001268; NID:G5302778; PIDN:CAB46056.1; GSPDB:GN00140
C:Genetics:
A:Gene: dl3551w
A:Map position: 4

Query Match 3.7%; Score 8; DB 2; Length 522;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 178 RQRKRHKHS 185
DB 69 RQRKRHKHS 76

RESULT 6
AF1490
hypothetical protein lin0462 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AF1490
R: Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, F
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunat, F.; Kurapkut, G.; Madueno, E.; Maitournam, A.; M
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1490
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <GLA>
A:CROSS-references: GB:AL592022; PIDN:CAC95694.1; PID:gl6412903; GSPDB:GN00178

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: lin0462

Query Match 3.2%; Score 7; DB 2; Length 100;

Best Local Similarity 100.0%; Pred. No. 20;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 SSDSGTS 120

Db 85 SSDSGTS 91

RESULT 7

AC1130

hypothetical protein lmo0442 [imported] - Listeria monocytogenes (strain EGD-e)

C;Species: Listeria monocytogenes

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C;Accession: AC1130

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kuraphat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Hierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AC1130

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-104 <GLA>

A;Cross-references: GB:NC_003210; PIDN:CA098521.1; PID:gl6409819; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo0442

Query Match 3.2%; Score 7; DB 2; Length 104;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 114 SSDSGTS 120

Db 88 SSDSGTS 94

RESULT 8

H97030

probable membrane protein [imported] - Clostridium acetobutylicum

C;Species: Clostridium acetobutylicum

C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C;Accession: H97030

R;Nolling, J.; Breton, G.; Onelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,

; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl

A;Reference number: A96900; MUID:21359325; PMID:21359325

A;Accession: H97030

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-105 <KUR>

A;Cross-references: GB:AE001437; PIDN:AAK79035.1; PID:gl5023973; GSPDB:GN00168

A;Experimental source: Clostridium acetobutylicum ATCC824

C;Genetics:

A;Gene: CAC1060

Query Match 3.2%; Score 7; DB 2; Length 105;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 201 IREICCE 207

Db 63 IREICCE 69

RESULT 9

T44366

hypothetical protein CzrA [imported] - Staphylococcus aureus

C;Species: Staphylococcus aureus

C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 11-May-2000

C;Accession: T44366

R;Kuroda, M.; Hayashi, H.; Ohta, T.

Microbiol. Immunol. 43, 115-125, 1999

A;Title: Chromosome-determined zinc responsible operon czr in Staphylococcus aureus str.

A;Reference number: Z22754; MUID:99244271; PMID:10229265

A;Accession: T44366

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-106 <KUR>

A;Cross-references: EMBL:AB016431; PIDN:BAA36687.1

A;Experimental source: strain 912

C;Genetics:

A;Note: CzrA

Query Match 3.2%; Score 7; DB 2; Length 106;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 LKLIKSV 41

Db 60 LKLIKSV 66

RESULT 10

S14234

hypothetical protein - mouse

C;Species: Mus musculus (house mouse)

C;Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 05-Nov-1999

C;Accession: S14234

R;Magyar, J.P.; Schachner, M.

Nucleic Acids Res. 18, 6695-6696, 1990

A;Title: Genomic structure of the adhesion molecule on glia (AMOG, Na/K-ATPase beta-2 s

A;Reference number: S14234; MUID:91067470; PMID:1701244

A;Accession: S14234

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 1-133 <WAG>

A;Cross-references: EMBL:X56007; NID:g50051; PIDN:CAA39481.1; PID:g50052

Query Match 3.2%; Score 7; DB 2; Length 133;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 32 PLLIKLL 38

Db 28 PLLIKLL 34

RESULT 11

H72476

hypothetical protein APE2457 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999

C;Accession: H72476

R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; I

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: H72476

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-133 <KAW>

A;Cross-references: DDBJ:AP000064; NID:g5105945; PIDN:BAA81472.1; PID:dl045258; PID:g51

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE2457

Query Match 3.2%; Score 7; DB 2; Length 133;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 83 DLFGVPS 89
DB 13 DLFGVPS 19
|||||

RESULT 12
JC5097
ferric uptake regulator - Haemophilus ducreyi
C:Species: Haemophilus ducreyi
C:Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 29-Sep-1999
C:Accession: JC5097
R:Carlson, S.D.B.; Thomas, C.E.; Elkins, C.
Gene 176, 125-129, 1996
A:Title: Cloning and sequencing of a Haemophilus ducreyi fur homolog.
A:Reference number: JC5097; MUID:97075920; PMID:8918243
A:Accession: JC5097
A:Molecule type: DNA
A:Residues: 1-149 <CAR>
A:Cross-references: GB:U37224; NID:G1568604; PIDN:AAC44583.1; PID:G1568605
C:Comment: This protein acts as a classical negative regulator, binding the conserved se
C:Genetics:
A:Gene: fur
C:Superfamily: ferric uptake regulator
C:Keywords: DNA binding; transcription regulation

Query Match 3.2%; Score 7; DB 2; Length 149;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 KLLKSVG 42
DB 7 KLLKSVG 13
|||||

RESULT 13
T10768
latex allergen Hev b5 - Para rubber tree
C:Species: Hevea brasiliensis (Para rubber tree)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: T10768
R:Slater, J.E.; Vedvick, T.; Arthur-Smith, A.; Trybul, D.E.; Kekwick, R.K.O.
J. Biol. Chem. 271, 25394-25399, 1996
A:Title: Identification, cloning and sequence of a major allergen (Hev b 5) from natural
A:Reference number: Z17129; MUID:96411753; PMID:8810306
A:Accession: T10768
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-151 <SLA>
A:Cross-references: EMBL:U42640; NID:G1480456; PIDN:AAC49447.1; PID:G1480457
A:Experimental source: cv. RRIM 600; latex

Query Match 3.2%; Score 7; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 PASQET 26
DB 36 PASQET 42
|||||

RESULT 14
E82330
conserved hypothetical protein VC0377 [imported] - Vibrio cholerae (strain N16961 serog
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: E82330
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bae, S.; Qin, H.; Dragoi, I.; Sellers, F

Query Match 3.2%; Score 7; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 153 SRPSTSS 159
DB 15 SRPSTSS 21
|||||

Search completed: March 14, 2004, 21:55:36
Job time : 22.6 secs

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: E82330
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-153 <HEI>
A:Cross-references: GB:AE004126; GB:AE003852; NID:G9654802; PIDN:AAF93550.1; GSPDB:GN00
A:Experimental source: serogroup O1, strain N16961, biotype El Tor
C:Genetics:
A:Gene: VC0377
A:Map position: 1

Query Match 3.2%; Score 7; DB 2; Length 153;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 192 FDESIAL 198
DB 61 FDESIAL 67
|||||

RESULT 15
T27416
hypothetical protein Y76A2A.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Sep-2000
C:Accession: T27416
R:Steward, C.
submitted to the EMBL Data Library, October 1998
A:Reference number: Z20362
A:Accession: T27416
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-187 <WIL>
A:Cross-references: EMBL:AL032665; PIDN:CAA21770.1; GSPDB:GN00021; CBSP:Y76A2A.3
A:Experimental source: clone Y76A2A
C:Genetics:
A:Gene: CBSP:Y76A2A.3
A:Map position: 3
A:Introns: 54/2; 118/1
C:Superfamily: heavy-metal-associated homology
P:53-82/Domain: heavy-metal-associated homology <HMA>

Query Match 3.2%; Score 7; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:52:59 ; Search time 32.7724 Seconds
(without alignments)
1391.692 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 216
Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIREICCRSSSESTG 216

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 809742 seqs, 211153259 residues

Word size : 0

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA.*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTU5_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	216	100.0	216	13	US-10-057-510-4
2	159	73.6	491	9	US-09-888-077-1
3	159	73.6	491	9	US-09-888-077-1
4	159	73.6	491	10	US-09-956-425-8
5	159	73.6	491	15	US-09-029-327-2
6	12	19.4	59	14	US-10-422-536-137
7	29	13.4	489	9	US-10-211-088-143
8	8	3.7	95	9	US-09-956-425-6
9	8	3.7	95	9	US-09-956-425-18
10	8	3.7	95	9	US-09-956-425-21
11	8	3.7	876	13	US-10-078-929-206
12	8	3.7	876	16	US-10-389-566-2043
13	8	3.7	876	16	US-10-389-566-2176
14	8	3.7	910	16	US-10-389-566-2431
15	7	3.2	20	9	US-09-865-553-2

16	7	3.2	49	9	US-09-864-761-41464	Sequence 41464, A
17	7	3.2	56	9	US-09-864-761-35082	Sequence 35082, A
18	7	3.2	95	9	US-09-956-425-19	Sequence 19, Appl
19	7	3.2	95	9	US-09-956-425-20	Sequence 20, Appl
20	7	3.2	142	14	US-10-156-761-10181	Sequence 10181, A
21	7	3.2	150	10	US-09-847-208-106	Sequence 106, App
22	7	3.2	160	9	US-09-798-042-106	Sequence 106, App
23	7	3.2	161	9	US-09-798-042-106	Sequence 106, App
24	7	3.2	188	16	US-10-389-566-588	Sequence 588, App
25	7	3.2	188	16	US-10-389-566-588	Sequence 588, App
26	7	3.2	261	9	US-09-159-469-72	Sequence 72, Appl
27	7	3.2	261	9	US-09-798-042-72	Sequence 72, Appl
28	7	3.2	264	15	US-10-369-493-20732	Sequence 20732, A
29	7	3.2	278	14	US-10-223-598-8	Sequence 8, Appl
30	7	3.2	323	9	US-09-159-469-9	Sequence 9, Appl
31	7	3.2	323	9	US-09-798-042-9	Sequence 9, Appl
32	7	3.2	326	14	US-10-146-772-4	Sequence 4, Appl
33	7	3.2	326	15	US-10-441-742-4	Sequence 4, Appl
34	7	3.2	326	15	US-10-440-523-4	Sequence 4, Appl
35	7	3.2	355	14	US-10-219-220-191	Sequence 191, App
36	7	3.2	391	14	US-10-223-598-6	Sequence 6, Appl
37	7	3.2	395	15	US-10-264-049-2980	Sequence 2980, Ap
38	7	3.2	440	15	US-10-369-493-17452	Sequence 17452, A
39	7	3.2	467	9	US-09-738-626-5847	Sequence 5847, Ap
40	7	3.2	502	14	US-10-157-031-92	Sequence 92, Appl
41	7	3.2	504	9	US-09-892-985-4	Sequence 4, Appl
42	7	3.2	504	14	US-10-349-836-4	Sequence 4, Appl
43	7	3.2	511	15	US-10-231-913-81	Sequence 81, Appl
44	7	3.2	511	15	US-10-231-913-82	Sequence 82, Appl
45	7	3.2	581	15	US-10-369-493-17186	Sequence 17186, A

ALIGNMENTS

RESULT 1

```
US-10-057-510-4
; Sequence 4, Application US/10057510
; Publication No. US20020098580A1
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDN INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT FILING DATE: 2002-01-25
; PRIOR FILING DATE: 2000-02-22
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-510-4
```

Query Match	100.0%	Score 216	DB 13	Length 216	
Best Local Similarity	100.0%	Pred. No. 2.3e-195			
Matches 216	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
Qy	1	MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYQYI	60		
Db	1	MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYQYI	60		
Qy	61	MTKELYDEKQOHVYCSNDLLGLDFGVPSVSKHKRIYMIYRNLYVNVQESDSDGTS	120		
Db	61	MTKELYDEKQOHVYCSNDLLGLDFGVPSVSKHKRIYMIYRNLYVNVQESDSDGTS	120		

QY 121 VSENCHLEGGSDQKDLVQELQEEKPSSSHLSVRPSTSSRRRAISETEENSDELSGER 180
DB 121 VSENCHLEGGSDQKDLVQELQEEKPSSSHLSVRPSTSSRRRAISETEENSDELSGER 180
QY 181 KRHKSDSISLSPDESALCVIREICCCERSSSESTG 216
DB 181 KRHKSDSISLSPDESALCVIREICCCERSSSESTG 216

RESULT 2
US-09-888-077-1
; Sequence 1, Application US/09888077
; Patent No. US20020031818A1
; GENERAL INFORMATION:
; APPLICANT: Ronai, Ze'ev
; APPLICANT: Fuchs, Serge
; TITLE OF INVENTION: Modification of Mdm2 Activity
; FILE REFERENCE: 2420/1H195-US1
; CURRENT APPLICATION NUMBER: US/09/888, 077
; CURRENT FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,343
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-077-1

Query Match 73.6%; Score 159; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.4e-141;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDGLFQVPFSVKEHKIYTMIRNLVVVNQESSDS 117
DB 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDGLFQVPFSVKEHKIYTMIRNLVVVNQESSDS 117
QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
DB 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
QY 178 RQRKHKSDSISLSPDESALCVIREICCCERSSSESTG 216
DB 178 RQRKHKSDSISLSPDESALCVIREICCCERSSSESTG 216

RESULT 3
US-09-956-425-8
; Sequence 8, Application US/09956425
; Patent No. US20020045192A1
; GENERAL INFORMATION:
; APPLICANT: Kriwacki, Richard
; APPLICANT: Bothner, Brian
; APPLICANT: Lewis, William
; TITLE OF INVENTION: Aif and Hdm2 Interaction Domains and Method of Use Thereof
; FILE REFERENCE: 1340/L/035
; CURRENT APPLICATION NUMBER: US/09/956,425
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 8
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-956-425-8

Query Match 73.6%; Score 159; DB 9; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.4e-141;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDGLFQVPFSVKEHKIYTMIRNLVVVNQESSDS 117

DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGLDGLFQVPFSVKEHKIYTMIRNLVVVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
DB 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKHKSDSISLSPDESALCVIREICCCERSSSESTG 216
DB 179 RQRKHKSDSISLSPDESALCVIREICCCERSSSESTG 217

RESULT 4
US-09-029-327-2
; Sequence 2, Application US/09029327
; Publication No. US20030060432A1
; GENERAL INFORMATION:
; APPLICANT: TOCQUE, Bruno
; APPLICANT: WASILYK, Bohdan
; APPLICANT: DUBS-POTERSZMAN,
; APPLICANT: Marie-Christine
; TITLE OF INVENTION: ANTAGONISTS OF THE ONCOGENIC ACTIVITY OF
; TITLE OF INVENTION: THE PROTEIN MDM2, AND USE THEREOF IN THE TREATMENT OF
; TITLE OF INVENTION: CANCERS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/029,327
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 96/01340
; FILING DATE: 02-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR95/10331
; FILING DATE: 04-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fehlner Esq., Paul F.
; REGISTRATION NUMBER: 35,135
; REFERENCE/DOCKET NUMBER: ST95050-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-029-327-2

Query Match 73.6%; Score 159; DB 10; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.4e-141;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDGLFQVPFSVKEHKIYTMIRNLVVVNQESSDS 117
DB 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDGLFQVPFSVKEHKIYTMIRNLVVVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
DB 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 178

QY 178 RQRKRKSDSISLFDLSALCVIREICCRSSSESTG 216
|
Db 179 RQRKRKSDSISLFDLSALCVIREICCRSSSESTG 217

RESULT 5
US-10-422-536-137
; Sequence 137, Application US/10422536
; Publication No. US20040014100A1
; GENERAL INFORMATION:
; APPLICANT: Kinsella, Todd
; APPLICANT: Lorens, James
; APPLICANT: Pray, Todd
; APPLICANT: Bennett, Mark
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR INHIBITING
; TITLE OF INVENTION: PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: A-71433-1/AMP/CVO
; CURRENT APPLICATION NUMBER: US/10/422,536
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: US 60/187,130
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: US 10/232,758
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 137
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-422-536-137

Query Match 73.6%; Score 159; DB 15; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.4e-141;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 QYIMTKELYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIYMIYRNLVVNNQESSDS 117
|
Db 59 QYIMTKELYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIYMIYRNLVVNNQESSDS 118
|
QY 118 GTSVSENRCHLEGGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISSEENSDLSGE 177
|
Db 119 GTSVSENRCHLEGGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISSEENSDLSGE 178
|
QY 178 RQRKRKSDSISLFDLSALCVIREICCRSSSESTG 216
|
Db 179 RQRKRKSDSISLFDLSALCVIREICCRSSSESTG 217

RESULT 6
US-10-211-088-143
; Sequence 143, Application US/10211088
; Publication No. US20030104479A1
; GENERAL INFORMATION:
; APPLICANT: Bright, Gary R.
; APPLICANT: Premkumar, D. David
; APPLICANT: Chen, Yih-Tai
; TITLE OF INVENTION: No. US20030104479A1el Fusion Proteins And Assays For Molecular Bi
; FILE REFERENCE: 01-1022-US
; CURRENT APPLICATION NUMBER: US/10/211,088
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/309,395
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/341,589
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 143
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Artificial Sequence

; FEATURE:
; OTHER INFORMATION: Binding domain
US-10-211-088-143

Query Match 19.4%; Score 42; DB 14; Length 59;
Best Local Similarity 100.0%; Pred. No. 6.8e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 QYIMTKELYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIY 99
|
Db 18 QYIMTKELYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIY 59
|

RESULT 7
US-09-956-425-6
; Sequence 6, Application US/09956425
; Patent No. US20020045192A1
; GENERAL INFORMATION:
; APPLICANT: Kriwacki, Richard
; APPLICANT: Bothner, Brian
; APPLICANT: Lewis, William
; TITLE OF INVENTION: Arf and Hdm2 Interaction Domains and Method of Use Thereof
; FILE REFERENCE: 1340/1/035
; CURRENT APPLICATION NUMBER: US/09/956,425
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-956-425-6

Query Match 13.4%; Score 29; DB 9; Length 489;
Best Local Similarity 100.0%; Pred. No. 7.7e-19;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 TSQIPASEQETLVPRKPLLLKLSVGAQ 44
|
Db 16 TSQIPASEQETLVPRKPLLLKLSVGAQ 44
|

RESULT 8
US-09-956-425-18
; Sequence 18, Application US/09956425
; Patent No. US20020045192A1
; GENERAL INFORMATION:
; APPLICANT: Kriwacki, Richard
; APPLICANT: Bothner, Brian
; APPLICANT: Lewis, William
; TITLE OF INVENTION: Arf and Hdm2 Interaction Domains and Method of Use Thereof
; FILE REFERENCE: 1340/1/035
; CURRENT APPLICATION NUMBER: US/09/956,425
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-956-425-18

Query Match 3.7%; Score 8; DB 9; Length 95;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 209 SSSSESTG 216
|
Db 1 SSSSESTG 8
|

RESULT 9
US-09-956-425-21

```
; Sequence 21, Application US/09956425
; Patent No. US20020045192A1
; GENERAL INFORMATION:
; APPLICANT: Kziwacki, Richard
; APPLICANT: Bothner, Brian
; APPLICANT: Lewis, William
; TITLE OF INVENTION: Arf and Hdm2 Interaction Domains and Method of Use Thereof
; FILE REFERENCE: 1340/1/035
; CURRENT APPLICATION NUMBER: US/09/956,425
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 95
; TYPE: PRT
; ORGANISM: horse
; US-09-956-425-21

Query Match      3.7%; Score 8; DB 9; Length 95;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      209 SSSSESTG 216
DB      1 SSSSESTG 8

RESULT 10
US-09-956-425-22
; Sequence 22, Application US/09956425
; Patent No. US20020045192A1
; GENERAL INFORMATION:
; APPLICANT: Kziwacki, Richard
; APPLICANT: Bothner, Brian
; APPLICANT: Lewis, William
; TITLE OF INVENTION: Arf and Hdm2 Interaction Domains and Method of Use Thereof
; FILE REFERENCE: 1340/1/035
; CURRENT APPLICATION NUMBER: US/09/956,425
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 95
; TYPE: PRT
; ORGANISM: dog
; US-09-956-425-22

Query Match      3.7%; Score 8; DB 9; Length 95;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      209 SSSSESTG 216
DB      1 SSSSESTG 8

RESULT 11
US-10-078-929-206
; Sequence 206, Application US/10078929
; Publication No. US20020152497A1
; GENERAL INFORMATION:
; APPLICANT: Rafaleki, Antoni
; APPLICANT: Miao, Guo-Hua
; APPLICANT: Falco, Saverio Carl
; APPLICANT: Sakai, Hajime
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Odell, Joan T.
; APPLICANT: Meyers, Blake
; APPLICANT: Thorpe, Catherine
; APPLICANT: Weng, Zude
; TITLE OF INVENTION: Nucleic Acid Fragments Encoding Proteins Involved in
; TITLE OF INVENTION: Stress Response
; FILE REFERENCE: BB1357 US NA
```

```
; CURRENT APPLICATION NUMBER: US/10/078,929
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/566,394
; PRIOR FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: 60/133038
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133042
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: 60/133427
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133437
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133428
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133438
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/133436
; PRIOR FILING DATE: 1999-05-11
; PRIOR APPLICATION NUMBER: 60/137667
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 206
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-10-078-929-206

Query Match      3.7%; Score 8; DB 13; Length 876;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      22 SEQETLVR 29
DB      471 SEQETLVR 478

RESULT 12
US-10-389-566-2043
; Sequence 2043, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2043
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-10-389-566-2043

Query Match      3.7%; Score 8; DB 16; Length 876;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      22 SEQETLVR 29
DB      471 SEQETLVR 478

RESULT 13
US-10-389-566-2176
```

; Sequence 2176, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2176
; LENGTH: 876
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-389-566-2176

Query Match 3.7%; Score 8; DB 16; Length 876;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 SEQETLVR 29
Db 471 SEQETLVR 478

RESULT 14
US-10-389-566-2431
; Sequence 2431, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2431
; LENGTH: 910
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-389-566-2431

Query Match 3.7%; Score 8; DB 16; Length 910;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 SEQETLVR 29
Db 459 SEQETLVR 466

RESULT 15
US-09-865-553-2
; Sequence 2, Application US/09865553
; Patent No. US20020055174A1
; GENERAL INFORMATION:
; APPLICANT: Rittner, Karola
; APPLICANT: Jacobs, Eric

; TITLE OF INVENTION: Complex for Transferring an Anionic Substance of Interest
; TITLE OF INVENTION: Into a Cell
; FILE REFERENCE: 032751-050
; CURRENT APPLICATION NUMBER: US/09/865,553
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/246,083
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 60/277,982
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: EP 00440162.6
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: EP 01440049.3
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PPTG1
US-09-865-553-2

Query Match 3.2%; Score 7; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 LLKLLKS 40
Db 6 LLKLLKS 12

Search completed: March 14, 2004, 21:57:32
Job time : 33.7724 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2004, 21:54:50 ; Search time 1111.79 Seconds
(without alignments)
2884.893 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 398

Sequence: 1 ARAYXKMFMEFKAGKWC.....ARLXPCGKVGMDVWRRWS 74

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US10057510/runat_09032004_162234_25975/app.query.fasta_1.654
-DB=GenEmbl -QFMT=fastap -SUFFIX=std.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=500 -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10057510@cgn_1.1.2471@runat_09032004_162234_25975 -NCPU=6 -ICPU=3
-NO_WMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	394	99.0	486	6	AR207093 Sequence
2	363	91.2	161179	2	AC083954 Homo sapi
C 3	363	91.2	161362	2	AL355503 Homo sapi
C 4	363	91.2	223735	9	AC096677 Homo sapi
5	165	41.5	1415	6	AX017292 Sequence
6	165	41.5	1415	6	AX524996 Sequence
7	165	41.5	1415	6	BD134463 Human nuc
8	165	41.5	1440	9	AF151100 Homo sapi
9	164	41.2	1076	6	BD192093 32 human
10	162	40.7	1384	9	BC014390 Homo sapi
11	157	39.4	1390	6	AR268917 Sequence
12	157	39.4	1390	6	BD218437 Molecules
13	154	38.7	1480	6	BD127556 Primer fo
14	154	38.7	1480	9	AK075179 Homo sapi
15	121	30.4	499	6	AR268920 Sequence
16	121	30.4	499	6	BD218456 Molecules
17	95	23.9	583	6	AR268921 Sequence
18	95	23.9	583	6	BD218457 Molecules
C 19	85.5	21.5	103067	2	AC097795 4 of
C 20	85.5	21.5	229055	2	AC133841 Rattus no
C 21	85.5	21.5	241633	2	AC096932 Rattus no
C 22	85	21.4	221424	2	AC129058 Rattus no
C 23	84	21.1	1326	6	AR390235 Sequence
C 24	84	21.1	2679	6	AR390219 Sequence
C 25	84	21.1	2813	1	AF016253 Klebsiell
C 26	82.5	20.7	137343	2	AC110173 Mus muscu
C 27	81.5	20.5	91830	9	AL365436 Human DNA
C 28	79	19.8	2175	5	BC047248 Xenopus lae
C 29	79	19.8	2941	5	U85969 Xenopus lae
C 30	78	19.6	4073	3	AY118647 Drosophill
C 31	78	19.6	85633	2	AC014423 Drosophill
C 32	78	19.6	159744	2	AC147424 Callicebu
C 33	78	19.6	169209	3	AC093499 Drosophill
C 34	78	19.6	202508	2	AC114610 Mus muscu
C 35	78	19.6	204744	2	AC122362 Mus muscu
C 36	78	19.6	252201	3	AE003529 Drosophill
C 37	78	19.6	266115	2	AC122201 Mus muscu
C 38	77.5	19.5	242980	2	AC146485 Silurana
C 39	77	19.3	149953	2	AP005000 Oryza sat
C 40	77	19.3	154494	2	AP005070 Oryza sat
C 41	77	19.3	185494	10	AC123823 Mus muscu
C 42	76.5	19.2	3864	3	AY122110 Drosophill
C 43	76.5	19.2	226554	2	AC103041 Rattus no
C 44	76	19.1	11945	5	AF405540 Gallus ga
C 45	76	19.1	12728	12	AF405538 Synthetic

ALIGNMENTS

RESULT 1

```
AR207093
LOCUS       AR207093               486 bp      DNA      Linear      PAT 20-JUN-2002
DEFINITION   Sequence 1 from patent US 6372490.
ACCESSION   AR207093
VERSION      AR207093.1  GI:21505894
KEYWORDS     Unknown.
SOURCE       Unknown.
ORGANISM     Unknown.
REFERENCE    1 (bases 1 to 486)
AUTHORS     Nandabalan,K., Yang,M. and Schulz,V.
TITLE       Nucleic acid encoding the MDM interacting protein
JOURNAL     Patent: US 6372490-A 1 16-APR-2002;
FEATURES    Location/Qualifiers
             source
               1..486
               /organism="unknown"
               /mol_type="unassigned DNA"

ORIGIN
-----
Alignment Scores:
Pred. No.:      4.06e-39      Length:      486
Score:          394.00      Matches:      74
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     98.99%      Indels:      0
DB:              6      Gaps:      0

US-10-057-510-2 (1-74) x AR207093 (1-486)

Qy  1  AlaArgAlaTyr***LysMetPheSerMetPheMetGluValGAlaClyLysCysTTPCys 20
Db  1  GCACGAGCGGATNAAAAAATGTTTCCATGTTATGGAAGAGCTGGGAGTGGCTGGT 60
Qy  21  LysMetProLysLeuIleAspThrProPheSerIleValAlaProAlaLeuThrAla 40
Db  61  AAAATGCCAAGCTCATATAGATACCTCTTCCATGTTGCCCCCTGCTAACTGCT 120
Qy  41  ValLeuSerCysGlnLeuArgCysSerLeuTTPLeuValGlyAlaArgLeu***ProCys 60
Db  121  GTTCTTTCTTGGCAGGTCGTTGTTCCCTCGCTGTGGGGCAGCGGTGTTCATGT 180
Qy  61  GlyLysValGluGlyMetAspValTTPArgArgArgTTPSer 74
Db  181  GGCAGGTGGAGGCATGACGTCGTGGAGGAGCGCTGGAGC 222

RESULT 2
AC083954
LOCUS       AC083954               161179 bp      DNA      linear      HTG 05-NOV-2000
DEFINITION   Homo sapiens chromosome UL clone RP11-790113, WORKING DRAFT
SEQUENCE, 15 unchromed pieces.
ACCESSION   AC083954
VERSION     AC083954.2  GI:11079632
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
             1 (bases 1 to 161179)
             Waterston,R.H.
             The sequence of Homo sapiens clone
             Unpublished
             2 (bases 1 to 161179)
             Waterston,R.H.
             Direct Submission
             Submitted (07-Oct-2000) Genome Sequencing Center, Washington
             University School of Medicine, 4444 Forest Park Parkway, St. Louis,
             MO 63108, USA
             On Nov 3, 2000 this sequence version replaced gi:10716661.

COMMENT
-----
Center: Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
```

```
----- Project Information -----
Center project name: H.NH0790113
----- Summary Statistics -----
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 150357 bases at least Q40
Consensus quality: 154392 bases at least Q30
Consensus quality: 156116 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 153779; sum-of-contigs
Quality coverage: 4.03 in Q20 bases; agarose-fp
Quality coverage: 4.49 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 6949: contig of 6949 bp in length
* 6950 7049: gap of unknown length
* 7050 14080: contig of 7031 bp in length
* 14081 14180: gap of unknown length
* 14181 23784: contig of 9604 bp in length
* 23785 23884: gap of unknown length
* 23885 31683: contig of 7799 bp in length
* 31684 42970: gap of unknown length
* 42971 43070: contig of 11187 bp in length
* 43071 53194: contig of 10124 bp in length
* 53195 53294: gap of unknown length
* 53295 71919: contig of 18625 bp in length
* 71920 72020: gap of unknown length
* 72020 92964: contig of 20945 bp in length
* 92965 93064: gap of unknown length
* 93065 139613: contig of 46549 bp in length
* 139614 139713: gap of unknown length
* 139714 141292: contig of 1579 bp in length
* 141293 141393: gap of unknown length
* 141393 143307: contig of 1915 bp in length
* 143308 143407: gap of unknown length
* 143408 146598: contig of 3191 bp in length
* 146599 149624: contig of 3026 bp in length
* 149625 149725: gap of unknown length
* 149725 154604: contig of 4880 bp in length
* 154605 154704: gap of unknown length
* 154705 161179: contig of 6475 bp in length.
*
* Location/Qualifiers
* 1..161179
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="UL"
* /clone="RP11-790113"
*
* 1..6949
* /note="assembly_name:Contig10"
* 7050..14080
* /note="assembly_name:Contig11"
* 14181..23784
* /note="assembly_name:Contig12"
* 23885..31683
* /note="assembly_name:Contig13"
* clone end.T7
* vector_side:left
* 31784..42970
* /note="assembly_name:Contig14"
* 43071..53194
* /note="assembly_name:Contig15"
```

```

misc_feature      /note="assembly_name:Contig15"
53295..71919
misc_feature      /note="assembly_name:Contig16"
72020..92964
misc_feature      /note="assembly_name:Contig17"
93065..139613
misc_feature      /note="assembly_name:Contig18"
139714..141292
misc_feature      /note="assembly_name:Contig19"
141393..143307
misc_feature      /note="assembly_name:Contig20"
143408..146598
misc_feature      /note="assembly_name:Contig21"
146699..149624
misc_feature      /note="assembly_name:Contig22"
149725..154604
misc_feature      /note="assembly_name:Contig23"
154705..161179
misc_feature      /note="assembly_name:Contig24"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      1,458-32      Length:      161179
Score:          363.00      Matches:      68
Percent Similarity: 97.14%      Conservative: 0
Best Local Similarity: 97.14%      Mismatches: 1
Query Match:      91.21%      Indels:      1
DB:              2          Gaps:      0

```

US-10-057-510-2 (1-74) x AC083954 (1-161179)

```

Qy      6  LysMetPheSerMetPheMetGlulysAlaGlyLysCysTrpCysLysMetProLysLeu 25
Db      1483 AAAATGTTTCCATGTTTATGAAAGCTGGGAAGTGTGCTGTAATAATGCCCAAGCTC 1542
Qy      26  IleIleAcTrpPropheSerIleValAlaProAlaLeuThrAlaValLeuSerCysGln 45
Db      1543 AATATAGATATCTCTTCCATGTTGCCCTGCTCTAATGCTGTCTTCTTTCGCAG 1602
Qy      46  LeuArgCysSerLeuTrpLeuValGlyAlaArgLeu...ProCysGlyLysValGluGl 65
Db      1603 CTTCGTTGTTCCCTGCTGCTGTTGGGGCACCGCTGTGCTCCATGTGGCAAGTGGAGG 1662
Qy      65  yMetAspValTrpArgArgTrpSer 74
Db      1663 CATGACGTGTGGAGGCGCTGGAGC 1690

```

RESULT 3

```

AL355503/c      AL355503      161362 bp      DNA      linear      HTG 10-JUL-2001
LOCUS           Homo sapiens chromosome 1 clone RP11-51H18, 20 unordered pieces.
DEFINITION      Homo sapiens chromosome 1 clone RP11-51H18, 20 unordered pieces.
ACCESSION       AL355503
VERSION         AL355503.3 GI:9797310
KEYWORDS        HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE           Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

```

1
AUTHORS         McIay,K.
TITLE           Direct Submission
JOURNAL         Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequest@sanger.ac.uk
On Aug 12, 2000 this sequence version replaced gi:9213570.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: b51H18
----- Summary Statistics

```

Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Consensus quality: 152267 bases at least Q40
Consensus quality: 155560 bases at least Q30
Consensus quality: 157235 bases at least Q20
Insert size: 159462; sum-of-contigs
Insert size: 173060; 1.1% error; agarose-fp
Quality coverage: 3.87x in Q20 bases; sum-of-contigs Quality
coverage: 3.62x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

```

*      1      8352: contig of 8352 bp in length
*      8353      8452: gap of 100 bp
*      8453      20258: contig of 11806 bp in length
*      20259      20358: gap of 100 bp
*      20359      24599: contig of 4601 bp in length
*      24599      25059: gap of 100 bp
*      25060      27559: contig of 2600 bp in length
*      27559      27759: gap of 100 bp
*      27760      30468: contig of 2709 bp in length
*      30469      30568: gap of 100 bp
*      30569      32558: contig of 2030 bp in length
*      32559      32758: gap of 100 bp
*      32759      34955: contig of 2197 bp in length
*      34956      35055: gap of 100 bp
*      35056      45995: contig of 10940 bp in length
*      45996      46095: gap of 100 bp
*      46096      57675: contig of 11580 bp in length
*      57676      57775: gap of 100 bp
*      57776      70772: contig of 12397 bp in length
*      70773      70872: gap of 100 bp
*      70873      88271: contig of 17399 bp in length
*      88272      88371: gap of 100 bp
*      88372      108997: contig of 20626 bp in length
*      108998      109097: gap of 100 bp
*      109098      119846: contig of 10749 bp in length
*      119847      119946: gap of 100 bp
*      119947      124556: contig of 4710 bp in length
*      124557      124756: gap of 100 bp
*      124757      129692: contig of 4936 bp in length
*      129693      129792: gap of 100 bp
*      129793      132724: contig of 2932 bp in length
*      132725      132824: gap of 100 bp
*      132825      144586: contig of 11762 bp in length
*      144587      144686: gap of 100 bp
*      144687      148461: contig of 3775 bp in length
*      148462      148561: gap of 100 bp
*      148562      155615: contig of 7054 bp in length
*      155616      155715: gap of 100 bp
*      155716      161362: contig of 5647 bp in length.

```

FEATURES

```

1..161362
source          /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="1"
                /clone="RP11-51H18"
                /clone_lib="RPC1-11.1"
                /note="assembly_fragment:00419
fragment_chain:1
clone_end:SF6
vector_side:left"
8453..20258
misc_feature    /note="assembly_fragment:01016"

```



```

misc_feature      20359..24959      fragment_chain:1"
                    /note="assembly fragment:00251"
misc_feature      25060..27659      fragment_chain:2"
                    /note="assembly fragment:00747"
misc_feature      27760..30468      fragment_chain:2"
                    /note="assembly fragment:00023"
misc_feature      30569..32658      fragment_chain:2"
                    /note="assembly fragment:00162"
misc_feature      32759..34955      fragment_chain:2"
                    /note="assembly fragment:00931"
misc_feature      35056..35995      fragment_chain:2"
                    /note="assembly fragment:00525"
misc_feature      40096..57675      fragment_chain:2"
                    /note="assembly fragment:01373"
misc_feature      57776..70772      fragment_chain:2"
                    /note="assembly fragment:00881"
misc_feature      70873..88271      fragment_chain:2"
                    /note="assembly fragment:00504"
misc_feature      89372..108997      fragment_chain:2"
                    /note="assembly fragment:00452"
misc_feature      109098..119846      fragment_chain:3"
                    /note="assembly fragment:00334"
misc_feature      119947..124656      fragment_chain:3"
                    /note="assembly fragment:00694"
misc_feature      124757..125692      fragment_chain:3"
                    /note="assembly fragment:00958"
misc_feature      129793..132724      fragment_chain:4"
                    /note="assembly fragment:00545"
misc_feature      132825..144586      fragment_chain:4"
                    /note="assembly fragment:00145"
misc_feature      144687..148461      fragment_chain:5"
                    /note="assembly fragment:01348"
misc_feature      148562..155615      fragment_chain:5"
                    /note="assembly fragment:01135"
misc_feature      155716..161362      fragment_chain:5"
                    /note="assembly fragment:00450"
                    clone_end:T7
                    vector_side:right"

ORIGIN
Alignment Scores:
Pred. No.:      1.45e-32      Length:      161362
Score:          363.00      Matches:      68
Percent Similarity: 97.14%      Conservative: 0
Best Local Similarity: 97.14%      Mismatches: 1
Query Match:    91.21%      Indels:      1
DB:             2      Gaps:      0

US-10-057-510-2 (1-74) x AL355503 (1-161362)

QY      6      LysMetPheSerMetPheMetGluLysValAlaGlyLysCysTrpCysLysMetProLysLeu 25
DB      6234      AAAATGTTTTCATGTTTATGAAAGGCTGGGAAGTGTGTTGTTAAATGCCAAGCTC 6175
QY      26      IleLeAspThrProPheSerIleValAlaProAlaLeuThraAlaValLeuSerCysGln 45

```

6174 ATAATAGTACTCTCTTCTCCATGTTGCCCTGCTTAAGTCTGTCTTCTTCCAG 6115
 46 LeuArgCysSerLeuTrpLeuValGlyAlaArgLeu-***ProCysGlyLysValGluG1 65
 6114 CTTCGTTGTTCCCTCTGCTTGTGGGGGACGGCTGTGCTCCATGTGCAAGGTGAAGG 6055
 65 YMetAspValTrpArgArgTrpSer 74
 6054 CATGACGTGTGGAGGAGGCTGGAGC 6027

RESULT 4
 AC096677/c
 LOCUS Homo sapiens chromosome 1 clone RP11-134G8, complete sequence. 223735 bp DNA linear PRI 02-APR-2002
 DEFINITION AC096677 AL157898
 ACCESSION AC096677.2 GI:19881834
 VERSION AC096677.2 GI:19881834
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 223735)
 Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
 Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
 Direct Submission
 Unpublished
 2 (bases 1 to 223735)
 Kaul,R.K., Olson,M.V., Raymond,C., Clendenning,J. and Haugen,E.D.
 Direct Submission
 Submitted (20-SEP-2001) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 3 (bases 1 to 223735)
 Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z.,
 Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
 Direct Submission
 Submitted (02-APR-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 On Apr 2, 2002 this sequence version replaced gi:15706065.
 ----- Genome Center
 Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: http://www.genome.washington.edu
 Contact: uwgctgs@u.washington.edu
 Drafting Center: SC
 ----- Project Information
 Center project name: chr-1
 Center clone name: RP11-134G8 (sc0639)
 ----- Summary Statistics
 Sequencing vector: plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator; ET; 89% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 223406 bases at least Q40
 Consensus quality: 223694 bases at least Q30
 Consensus quality: 223735 bases at least Q20
 Insert size: 223735; sum-of-contigs
 Quality coverage: 7.5x in Q20 bases; sum-of-contigs

Overlapping Sequences:
 5': RP11-27G14 (UWGC:sc0604) AL390790
 3': RP11-275J16 AL512788, 2000-bp overlap

 Sequence Quality Assessment:
 This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:
all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

HindIII			BglII			EcoRI		
SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap
3729	3708	3053	2979	8696	8763			
6382	6744	2067	2031	24	<800			
512	<800	11775	11600	291	<800			
449	<800	2644	2607	21165	21051			
1927	2048	2546	2607	13190	12847			
910	908	2627	2607	15596	15250			
1135	1067	3879	4036	5152	5143			
9758	9631	7627	7639	1918	1906			
2241	2244	3671	3824	9537	9534			
2607	2623	4048	4036	2334	2289			
9951	9631	3716	3654	8224	8256			
2865	2846	931	914	15735	15250			
1008	1067	2596	2607	1156	1137			
246	<800	3904	4036	2564	2578			
6892	6744	2074	2031	2881	2885			
2119	2048	2651	2607	2458	2578			
1513	1425	3662	3654	6232	6207			
13006	12636	3528	3484	419	<800			
3884	3871	4281	4552	1017	1001			
2031	2048	236	<800	2566	2578			
927	908	2363	2383	19096	19031			
1113	1067	11447	11163	6118	6207			
761	786	3265	3202	26436	26981			
2381	2390	7011	6948	1592	1553			

1387	1425	2159	2137	5776	5743
2570	2623	15	<800	4374	4372
10872	10635	335	<800	10701	10558
1787	1790	8090	8058	2886	2885
1455	1425	1460	1410	8185	8256
5126	5111	6256	6192	4595	4596
2416	2390	4697	4552	3674	3715
1234	1206	1846	1814	4567	4596
7407	7405	190	<800	2994	3014
6624	6744	4556	4552	1394	1360
570	<800	1047	1026	6838	6900
1771	1790	5744	5760	274	<800
3169	3134	9864	9854	1807	1769
16293	16547	1453	1410		
808	786	483	<800		
1091	1067	6319	6192		
4510	4463	4066	4235		
10976	10635	2964	2979		
148	<800	3022	2979		
3115	3134	382	<800		
283	<800	2131	2137		
14509	14410	14825	14670		
4740	4648	8658	8664		
1808	1790	1856	1901		
6913	6744	1682	1632		
466	<800	1932	2031		
1016	1067	5692	5601		
1056	1067	6615	6585		
2811	2846	5136	5117		
6185	6097	20043	20389		
180	<800	2554	2607		
2622	2623	778	774		
15457	15534				
83	<800				

Alignment Scores:

Pred. No.: 2,078-32
Score: 383.00
Percent Similarity: 97.14%
Best Local Similarity: 97.14%

Length: 223735
Matches: 68
Conservative: 0
Mismatch: 1

```

Query Match: 91.21% Indels: 1
DB: 9 Gaps: 0
US-10-057-510-2 (1-74) x AC096677 (1-223735)

Qy 6 LysMetPheSerMetPheMetGluLysAlaGlyLysCysTrpCysLysMetProLysLeu 25
Db 64296 AAAATGTTTTTCCATGTTTATGAAAGGCTGGGAAGTGTGCTGTAAATGCCAAGCTC 64237

Qy 26 IleileAspThrProPheSerIleValAlaProAlaLeuThrAlaValLeuSerCysGln 45
Db 64236 ATAATAGATACCTCTTCCATGTTGCGCTGTCTTAACGCTGTTCTTCTTGCCAG 64177

Qy 46 LeuArgCysSerLeuTrpLeuValGlyAlaArgLeu-***ProCysGlyLysValGluG1 65
Db 64176 CTTGCTGTGTTCCCTCTGCTGTGGGGCAGCGCTGTGCTCCATGTGGCAAGTGAAGG 64117

Qy 65 yMetAspValTrpArgArgTrpSer 74
Db 64116 CATGGACGTGTGGAGGCGCTGGAGC 64089

RESULT 5
LOCUS AX017292 1415 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 47 from Patent WO9947669.
ACCESSION AX017292
VERSION AX017292.1 GI:10042210
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Schmitt,A., Specht,T., Dahl,E., Hinzmann,B., Rosenthal,A. and
Pilarczyk,C.
TITLE Human nucleic acid sequences from tissue of breast tumors
JOURNAL PATENT: WO 9947669-A 47 23-SEP-1999;
SCHMITT THOMAS (DE); SPECHT THOMAS (DE); DAHL EDGAR (DE); HINZMANN
BERND (DE); ROSENTHAL ANDRE (DE); METAGEN GES FUER GENOMFORSCHUN
(DE); PILARSKY CHRISTIAN (DE)
FEATURES
source
1..1415
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores: 1.74e-10 Length: 1415
Pred. No.: 165.00 Matches: 32
Score: 79.55% Conservative: 3
Percent Similarity: 72.73% Mismatches: 8
Best Local Similarity: 41.46% Indels: 1
Query Match: 6 Gaps: 0
DB:

US-10-057-510-2 (1-74) x AX524996 (1-1415)
Qy 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 686 TCCTTCATGCTACCCACCTCAGTGTGCTGAGGTCAAGGAGCTTCGTGTTCCTCTGG 745

Qy 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 746 CTTGTGGGGGACCGCTGTGCTCCATGTGGCAAGTGAAGGATGGACGTGTGGAGGAG 805

Qy 71 gArgTrpSer 74
Db 806 GCGCTGGAGC 815

RESULT 7
LOCUS BD134463 1415 bp DNA linear PAT 18-SEP-2002
DEFINITION Human nucleic acid sequence originating in mammary tumor tissue.
ACCESSION BD134463
VERSION BD134463.1 GI:23229408
KEYWORDS JP 2002506643-A/41.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Speft,T., Hintzman,B., Armin,S., Pirarski,C., Edgar,D. and
Rosenthal,A.
TITLE Human nucleic acid sequence originating in mammary tumor tissue
JOURNAL Patent: JP 2002506643-A 41 05-MAR-2002;
METAGEN GESELLSCHAFT FUER GENOME FORSCHUNG MBH
COMMENT OS Homo sapiens (human)
PN JP 2002506643-A/41
PD 05-MAR-2002
PF 19-MAR-1999 JP 2000536852
PR 20-MAR-1998 DE 198 13 839.3
PI THOMAS SPEFT,BERND HINTZMAN, SHCMITT ARMIN,CHRISTIAN PIRARSKI,
PI DUHL EDGAR,
PI ANDRE ROSENTHAL
PC C12N15/09,A61K31/711,A61K38/00,A61K48/00,A61P35/00,C07K14/47,
PC C07K16/18,
PC C12N1/19,C12N5/10,C12N15/00,A61K37/02,C12N5/00 CC Human
nucleic acid sequence originating in mammary tumor CC
tissue
FH Key Location/Qualifiers

```


AR268917
LOCUS AR268917 1390 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 2 from patent US 6500642.
ACCESSION AR268917
VERSION AR268917.1 GI:29699652
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1390)
AUTHORS Yue,H., Patterson,C., Corley,N.C. and Guegler,K.J.
TITLE Molecule associated with apoptosis
JOURNAL Patent: US 6500642-A 2 31-DEC-2002;
FEATURES
Location/Qualifiers
1..1390
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 1.64e-09 Length: 1390
Score: 157.00 Matches: 31
Percent Similarity: 77.27% Conservative: 3
Best Local Similarity: 70.45% Mismatches: 9
Query Match: 39.45% Indels: 1
DB: 6 Gaps: 0
US-10-057-510-2 (1-74) x AR268917 (1-1390)
QY 32 SerIleValAlaProAlaLeuThraValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 655 TCCTTCATGCTACCCACACCTCAGTGTGAGGTCAAGGCAGCTTTGTGTCCCTCTGG 714
QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 715 CTTGTGGGGACGGCTGTCTCCATGTGGCAAGGTGGAGGCATGACGTGTGGAGGAG 774
QY 71 gArgTrpSer 74
Db 775 GCGCTGGAGC 784
RESULT 12
LOCUS BD218437 1390 bp DNA linear PAT 17-JUL-2003
DEFINITION Molecules associated with apoptosis.
ACCESSION BD218437
VERSION BD218437.1 GI:33028207
KEYWORDS JP 2002519029-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillman,J.L., Yue,H., Tang,Y.T., Corley,N.C., Guegler,K.J. and
Patterson,C.
TITLE Molecules associated with apoptosis
JOURNAL Patent: JP 2002519029-A 3 02-JUL-2002;
COMMENT INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002519029-A/3
PD 02-JUL-2002
PF 23-JUN-1999 JP 2000557362
PR 29-JUN-1998 US 09/106920
PI JENNIFER L HILLMAN,HENRY YUE,Y TOM TANG,NEIL C CORLEY,KARL J
PI GUEGLER,
PI CHANDRA PATTERSON
PI C12N15/09,A61K38/00,A61K45/00,A61P37/02,A61P43/00,
PC C07K14/47,
PC C07K14/705,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/10, PC
C12P2/02,
PC C12Q1/68,C12N15/00,A61K37/02,C12N5/00
CC Incyte clone 2840978
FH Key Location/Qualifiers

FT source 1..1390
/organism="Homo sapiens (human)".
FEATURES
Location/Qualifiers
1..1390
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.64e-09 Length: 1390
Score: 157.00 Matches: 31
Percent Similarity: 77.27% Conservative: 3
Best Local Similarity: 70.45% Mismatches: 9
Query Match: 39.45% Indels: 1
DB: 6 Gaps: 0
US-10-057-510-2 (1-74) x BD218437 (1-1390)
QY 32 SerIleValAlaProAlaLeuThraValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 655 TCCTTCATGCTACCCACACCTCAGTGTGAGGTCAAGGCAGCTTTGTGTCCCTCTGG 714
QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 715 CTTGTGGGGACGGCTGTCTCCATGTGGCAAGGTGGAGGCATGACGTGTGGAGGAG 774
QY 71 gArgTrpSer 74
Db 775 GCGCTGGAGC 784
RESULT 13
LOCUS BD127556 1480 bp DNA linear PAT 18-SEP-2002
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD127556
VERSION BD127556.1 GI:23222501
KEYWORDS JP 2002017375-A/2987.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ota,T., Nishikawa,T., Isogai,T., Hayashi,K., Ishii,S., Kawai,Y.,
Wakamatsu,A., Sugiyama,T., Nagai,K., Kojima,S., Otsuki,T. and
Koga,H.
TITLE Primer for synthesizing full-length cDNA and use thereof
JOURNAL Patent: JP 2002017375-A 2987 22-JAN-2002;
COMMENT HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002017375-A/2987
PD 22-JAN-2002
PF 07-JUL-2000 JP 2000253172
PI TOSHIO OTA,TETSUO NISHIKAWA,TAKAO ISOGAI,KOJI HAYASHI,SHIZUKO
PI ISHII,
PI YURI KAWAI,AI WAKAMATSU,TOMOYASU SUGIYAMA,KEIICHI NAGAI, PI
SHINICHI KOJIMA,
PI TETSUJI OTSUKI,HISASHI KOGA
PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT CDS Location/Qualifiers
(349)..(729).
1..1480
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:

```
Pred. No.: 4,11e-09 Length: 1480
Score: 154.00 Matches: 30
Percent Similarity: 77.2% Conservative: 4
Best Local Similarity: 68.18% Mismatches: 9
Query Match: 38.69% Indels: 1
DB: 6 Gaps: 0

US-10-057-510-2 (1-74) x BD127556 (1-1480)

QY 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 753 TCCTTCATGCTACCCACCACTTCATGCTGAGGACAGGACGCTTTGTTGCTCCCTCTGG 812
QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 813 CTGTGGGGGACCGGCTGTGCTCCATGTGGCAAGGTGGAAGGCATGGACGTGTGGAGGAG 872
QY 71 gArgTrpSer 74
Db 873 GCGCTGGAGC 882

RESULT 14
AK075179 1480 bp mRNA linear PRI 03-SEP-2002
LOCUS Homo sapiens cDNA FLJ90698 fis, clone PLACE1007028.
DEFINITION AK075179
ACCESSION AK075179.1 GI:22761100
VERSION oligo capping; fis (full insert sequence).
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Isogai,T., Ota,T., Nishikawa,T., Hayashi,K., Otsuki,T.,
Sugiyama,T., Suzuki,Y., Nagai,K., Sugano,S., Ishii,S.,
Kawai-Hio,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y.,
Kojima,S., Nagahara,K., Masuho,Y., Ono,T., Okano,K., Yoshikawa,Y.,
Aotsuka,S., Sasaki,N., Hattori,A., Okumura,K., Iwayanagi,T. and
Ninomiya,K.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1480)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (25-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology; cDNA library construction:
Institute of Medical Science, University of Tokyo; Laboratory of
Genome Structure, Human Genome Center; cDNA 5'- & 3'-end one pass
sequencing and clone selection: Helix Research Institute (supported
by Japan Key Technology Center etc.).
Location/Qualifiers
FEATURES
source
1..1480
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1007028"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="cloning vector: pME18SFL3"
349..732
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC11454.1"
/db_xref="GI:22761101"
/translation="MTAAATATLVKEGVLEKRSGLLQWLKRCVLTERRGLQLFEAK
GTGGRPKLSFARIKAVCVSTGRHIVFTLVTEGGGIDFRCLLEDPGWNAQITLGL
VKFKNQAIQIVRAQSLGTGLNS"

ORIGIN

Alignment Scores:
Pred. No.: 1,42e-05 Length: 499
Score: 121.00 Matches: 26
Percent Similarity: 69.05% Conservative: 3
Best Local Similarity: 61.90% Mismatches: 13
Query Match: 30.40% Indels: 1
DB: 6 Gaps: 0

US-10-057-510-2 (1-74) x AR268920 (1-499)

QY 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 354 TCCTTCATGCTACCCACCACTTCATGCTGAGGACAGGACGCTTCGTTGCCNCTGG 413
QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 414 CTGTGGGGGACCGGCTGTGCTCCATGTGGAGGAGTGAAGGCAT-CACGTGTGGANGANG 472
QY 72 ArgTrp 73
Db 473 GCGTGG 478

Search completed: March 14, 2004, 23:19:48
Job time : 1170.29 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:45:03 ; Search time 57.3517 Seconds
(without alignments)
795.253 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 1095

Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIRICRSGSSSSSTG 216

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/FCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1095	100.0	216	13	US-10-057-510-4
2	1084.5	99.0	491	9	US-09-888-077-1
3	1084.5	99.0	491	9	US-09-956-425-8
4	1084.5	99.0	491	10	US-09-029-327-2
5	1084.5	99.0	491	15	US-10-422-536-137
6	804.5	73.5	489	9	US-09-956-425-6
7	297.5	27.2	59	14	US-10-211-088-143
8	100	9.1	2184	14	US-10-304-095-6
9	93	8.5	485	11	US-09-833-245-1730
10	93	8.5	485	11	US-09-833-245-1731
11	93	8.5	485	11	US-09-833-245-1732
12	92.5	8.4	433	8	US-08-945-038-6
13	89.5	8.2	963	14	US-10-078-531-5
14	89	8.1	951	14	US-10-078-531-3
15	89	8.1	969	14	US-10-078-531-8

16	89	8.1	971	14	US-10-078-531-7	Sequence 7, Appli
17	89	8.1	1008	14	US-10-078-531-2	Sequence 2, Appli
18	88.5	8.1	1215	9	US-09-775-181-2	Sequence 2, Appli
19	88.5	8.1	1215	14	US-10-233-539-2	Sequence 2, Appli
20	88	8.0	438	14	US-10-176-584A-2	Sequence 4064, Ap
21	88	8.0	1170	15	US-10-369-493-4064	Sequence 3799, Ap
22	87.5	8.0	441	14	US-10-043-487-369	Sequence 369, Ap
23	87.5	8.0	481	15	US-10-104-047-3799	Sequence 4603, Ap
24	87.5	8.0	431	15	US-10-108-260A-4603	Sequence 46, Appl
25	87.5	8.0	515	10	US-09-315-355-46	Sequence 25, Appl
26	87.5	8.0	545	10	US-09-849-602-25	Sequence 105, App
27	87.5	8.0	572	10	US-09-738-630-105	Sequence 2460, Ap
28	87.5	8.0	1047	15	US-10-369-493-2460	Sequence 3, Appli
29	87	7.9	1781	9	US-09-738-877-3	Sequence 13, Appl
30	87	7.9	1781	10	US-09-961-403-13	Sequence 16, Appl
31	87	7.9	1781	15	US-10-428-487-16	Sequence 3133, Ap
32	86.5	7.9	754	15	US-10-094-749-3133	Sequence 78, Appl
33	86.5	7.9	1417	9	US-09-753-143-78	Sequence 859, App
34	86.5	7.9	1417	15	US-10-374-077-78	Sequence 8444, Ap
35	86	7.9	758	9	US-09-925-299-859	Sequence 62, Appl
36	86	7.9	758	14	US-09-925-299-859	Sequence 31, Appl
37	85	7.8	860	14	US-10-128-714-8444	Sequence 5, Appli
38	85	7.8	3051	15	US-10-144-194A-62	Sequence 21, Appl
39	84.5	7.7	223	15	US-10-120-835-31	Sequence 80, Appl
40	84.5	7.7	223	15	US-10-120-835-32	Sequence 82, Appl
41	84.5	7.7	295	15	US-10-120-835-5	Sequence 37, Appl
42	84.5	7.7	295	15	US-10-120-835-21	
43	84.5	7.7	295	15	US-10-120-835-80	
44	84.5	7.7	295	15	US-10-120-835-82	
45	84.5	7.7	310	15	US-10-120-835-37	

ALIGNMENTS

RESULT 1

US-10-057-510-4
; Sequence 4, Application US/10057510
; Publication No. US20020098580A1
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDM INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT APPLICATION NUMBER: US/10/057,510
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: USSN 09/510,252
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USSN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USSN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-510-4

Query Match 100.0%; Score 1095; DB 13; Length 216;
Best Local Similarity 100.0%; Pred.No. 2.1e-98;
Matches 216; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLKLLKSVGAQKDTYTMKEVFLQYI 60

Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLKLLKSVGAQKDTYTMKEVFLQYI 60

Qy 61 MTKRLYDEKQOHHIVYCSNDLLGLDPVPSFVKEHKKIYTMIRNLVYVNOQSSSGTS 120

Db 61 MTKRLYDEKQOHHIVYCSNDLLGLDPVPSFVKEHKKIYTMIRNLVYVNOQSSSGTS 120

Qy 121 VSENCHLEGSDQKDLVQLOEKPSSSHLSVRPSTSSRRRAISRETEENSDELGERQ 180
Db 121 VSENCHLEGSDQKDLVQLOEKPSSSHLSVRPSTSSRRRAISRETEENSDELGERQ 180
Qy 181 KRHKSDSISLSPDESALCVIREICCRSSSSSESTG 216
Db 181 KRHKSDSISLSPDESALCVIREICCRSSSSSESTG 216

RESULT 2

US-09-888-077-1
; Sequence 1, Application US/09888077
; Patent No. US2002003181A1
; GENERAL INFORMATION:
; APPLICANT: Ronai, Ze'ev
; APPLICANT: Fuchs, Serge
; TITLE OF INVENTION: Modification of Mdm2 Activity
; FILE REFERENCE: 2420/1H195-US1
; CURRENT APPLICATION NUMBER: US/09/888,077
; PRIOR FILING DATE: 2001-06-22
; PRIOR APPLICATION NUMBER: US 60/213,343
; PRIOR FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-077-1

Query Match 99.0%; Score 1084.5; DB 9; Length 491;
Best Local Similarity 99.5%; Pred. No. 6.8e-97;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 MCNTNMSVPTDGAVTTSQIPASEQETLVVRPKPILLKLLKSVGAQKDTYTMKEVLFYL-QY 59
Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVVRPKPILLKLLKSVGAQKDTYTMKEVLFYLGQY 60
Qy 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPSVKEHRKIYTMIRNLVWVNNQESSDSGT 119
Db 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPSVKEHRKIYTMIRNLVWVNNQESSDSGT 120
Qy 120 SVSENCHLEGSDQKDLVQLOEKPSSSHLSVRPSTSSRRRAISRETEENSDELGERQ 179
Db 121 SVSENCHLEGSDQKDLVQLOEKPSSSHLSVRPSTSSRRRAISRETEENSDELGERQ 180
Qy 180 KRHKSDSISLSPDESALCVIREICCRSSSSSESTG 216
Db 181 KRHKSDSISLSPDESALCVIREICCRSSSSSESTG 217

RESULT 3

US-09-956-425-8
; Sequence 8, Application US/09956425
; Patent No. US20020045192A1
; GENERAL INFORMATION:
; APPLICANT: Kriwacki, Richard
; APPLICANT: Bothner, Brian
; APPLICANT: Lewis, William
; TITLE OF INVENTION: Arf and Hdm2 Interaction Domains and Method of Use Thereof
; FILE REFERENCE: 1340/1/035
; CURRENT APPLICATION NUMBER: US/09/956,425
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-956-425-8

Query Match 99.0%; Score 1084.5; DB 9; Length 491;

Best Local Similarity 99.5%; Pred. No. 6.8e-97;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1 MCNTNMSVPTDGAVTTSQIPASEQETLVVRPKPILLKLLKSVGAQKDTYTMKEVLFYL-QY 59
Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVVRPKPILLKLLKSVGAQKDTYTMKEVLFYLGQY 60
Qy 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPSVKEHRKIYTMIRNLVWVNNQESSDSGT 119
Db 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPSVKEHRKIYTMIRNLVWVNNQESSDSGT 120
Qy 120 SVSENCHLEGSDQKDLVQLOEKPSSSHLSVRPSTSSRRRAISRETEENSDELGERQ 179
Db 121 SVSENCHLEGSDQKDLVQLOEKPSSSHLSVRPSTSSRRRAISRETEENSDELGERQ 180
Qy 180 KRHKSDSISLSPDESALCVIREICCRSSSSSESTG 216
Db 181 KRHKSDSISLSPDESALCVIREICCRSSSSSESTG 217

RESULT 4

US-09-029-327-2
; Sequence 2, Application US/09029327
; Publication No. US20030060432A1
; GENERAL INFORMATION:
; APPLICANT: TOCQUE, Bruno
; APPLICANT: WASLYK, Bohdan
; APPLICANT: DUBS-POTERSZMAN,
; APPLICANT: Marie-Christine
; TITLE OF INVENTION: ANTAGONISTS OF THE ONCOGENIC ACTIVITY OF
; TITLE OF INVENTION: THE PROTEIN MDM2, AND USE THEREOF IN THE TREATMENT OF
; TITLE OF INVENTION: CANCERS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3C43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/029,327
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 96/01340
; FILING DATE: 02-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR95/10331
; FILING DATE: 04-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fehlnher Esq., Paul F.
; REGISTRATION NUMBER: 35,135
; REFERENCE/DOCKET NUMBER: ST95050-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-029-327-2

Query Match 99.0%; Score 1084.5; DB 10; Length 491;
Best Local Similarity 99.5%; Pred. No. 6.8e-97;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVKPKLLKLLKSVGAOKDTYTMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVKPKLLKLLKSVGAOKDTYTMKEVLFYLGQY 60
QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 120
QY 120 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGERQ 179
DB 121 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGERQ 180
QY 180 RKRHKSISLSPDESALCVIREICCRSSSSSESTG 216
DB 181 RKRHKSISLSPDESALCVIREICCRSSSSSESTG 217

RESULT 5

US-10-422-536-137
; Sequence 137, Application US/10422536
; Publication No. US20040014100A1
; GENERAL INFORMATION:
; APPLICANT: Kinsella, Todd
; APPLICANT: Lorens, James
; APPLICANT: Pray, Todd
; APPLICANT: Bennett, Mark
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR INHIBITING
; TITLE OF INVENTION: PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: A-71433-1/AMP/CYO
; CURRENT APPLICATION NUMBER: US/10/422,536
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: US 60/187,130
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: US 10/232,758
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 137
; TYPE: PRT
; LENGTH: 491
; ORGANISM: Homo sapiens
US-10-422-536-137

Query Match 99.0%; Score 1084.5; DB 15; Length 491;
Best Local Similarity 99.5%; Pred. No. 6.8e-97;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVKPKLLKLLKSVGAOKDTYTMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVKPKLLKLLKSVGAOKDTYTMKEVLFYLGQY 60
QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 120
QY 120 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGERQ 179
DB 121 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGERQ 180
QY 180 RKRHKSISLSPDESALCVIREICCRSSSSSESTG 216
DB 181 RKRHKSISLSPDESALCVIREICCRSSSSSESTG 217

RESULT 6

US-09-956-425-6
; Sequence 6, Application US/09956425
; Patent No. US20020045192A1
; GENERAL INFORMATION:
; APPLICANT: Kziwacki, Richard

; APPLICANT: Bothner, Brian
; APPLICANT: Lewis, William
; TITLE OF INVENTION: Arf and Hdm2 Interaction Domains and Method of Use Thereof
; FILE REFERENCE: 1340/1/035
; CURRENT APPLICATION NUMBER: US/09/956,425
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 489
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-956-425-6

Query Match 73.5%; Score 804.5; DB 9; Length 489;
Best Local Similarity 76.0%; Pred. No. 1.2e-69;
Matches 168; Conservative 16; Mismatches 24; Indels 13; Gaps 4;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVKPKLLKLLKSVGAOKDTYTMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVKPKLLKLLKSVGAOKDTYTMKEVLFYIGQY 60
QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 117
QY 120 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGERQ 179
DB 118 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGERH 177
QY 180 RKRHKSISLSPDESALCVIREICCRSSSSSEST 215
DB 178 RKRHKSISLSPDESALCVIREICCRSSSSSEST 214

RESULT 7

US-10-211-088-143
; Sequence 143, Application US/10211088
; Publication No. US20030104479A1
; GENERAL INFORMATION:
; APPLICANT: Bright, Gary R.
; APPLICANT: Premkumar, D. David
; APPLICANT: Chen, Yih-Tai
; TITLE OF INVENTION: No. US20030104479A1el Fusion Proteins And Assays For Molecular B:
; FILE REFERENCE: 01-1022-US
; CURRENT APPLICATION NUMBER: US/10/211,088
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/309,395
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/341,589
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 143
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Binding domain
US-10-211-088-143

Query Match 27.2%; Score 297.5; DB 14; Length 59;
Best Local Similarity 98.3%; Pred. No. 1.3e-21;
Matches 58; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 42 GAOKDTYTMKEVLFYL-QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIY 99
DB 1 GAOKDTYTMKEVLFYLGQYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIY 59

RESULT 8

US-10-304-095-6
; Sequence 6, Application US/10304095

```

; Publication No. US20030134275A1
; GENERAL INFORMATION:
; APPLICANT: Long, David M.
; APPLICANT: Metz, Anneke M.
; APPLICANT: Love, Ruschelle A.
; TITLE OF INVENTION: Telomerase Reverse Transcriptase (TERT) Genes
; FILE REFERENCE: 47714-5009-US
; CURRENT APPLICATION NUMBER: US/10/304,095
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: US/09/417,485
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2184
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (330)..(335)
; OTHER INFORMATION: Xaa at position 330 = Leu or Ile;
; OTHER INFORMATION: Xaa at position 335 = Asp or Gly.
; US-10-304-095-6

Query Match      9.1%; Score 100; DB 14; Length 2184;
Best Local Similarity 23.7%; Pred. No. 3.3;
Matches 50; Conservative 33; Mismatches 82; Indels 46; Gaps 11;

QY      1 MCNTNMSVPTDG-----AVTTSQIPASEQETLVPRKPLLKLLKSVGAQKDTYTMKEVLFY 56
DB      139 VCTTKNISNDTSDKCTITKNIP-----LKYHINKKYKYLKK-KYH 180
QY      57 LQYIMTKRLYDEKQCHIVYCSNDLLGLFGVPSVKEHKIYTMIRNVLVAVN 110
DB      181 TWYTNNDHSYG-KYLVLVQCSGRLLKNDFFKDMKQIQEERKXYTSNIKINSEYTNIIIN 239
QY      111 QOESSDSGTSVSNRCHLEGSDQKLVQELQEKPS--SHLVSRPSTSSRRRAISETE 168
DB      240 NNNNNNNNNNNNNVH--GFGHNNLFS--SNEFPSSNISCTNYTEKNDKLTHIRETS 295
QY      169 ----ENS--DELSGE---RQRKRKSDSI 188
DB      296 LLITENSKKDKLLPEIDPFSEDEKESSSV 326

RESULT 9
US-09-833-245-1730
; Sequence 1730, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1730
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-833-245-1731

Query Match      8.5%; Score 93; DB 11; Length 485;
Best Local Similarity 21.9%; Pred. No. 1.9;
Matches 51; Conservative 39; Mismatches 77; Indels 66; Gaps 12;

QY      22 SEQETLVPRKP---LLLKLKSVGAQKDTYTMKEVLFYQYIMTKRLYDEKQCHIVYCSN 78
DB      259 SDPELGVBEPPHTIMKQLLSYITKDKQTESLVEKL--CQRFRTSRT-ERQQRDLAYCVS 315
QY      79 -----DILGLDFGVPS--FSVKEHKIYTMIRNVLVAVNQQESSDSGTSV 121
DB      316 QLPLTERGLRKLNDNFDFCGDKLSDSIFSA-----FLSVGKLRGAKPEGKAI 365
QY      122 SE-----NRCHLEG--GSDQKDLVQELQEKPS--SHLVSRPSTSSRRRAISETEENSDE 173
DB      366 IDEFEQKLKACHTRGLDGIKELEIGQAGSQAPSA---KKPSTGSRVQPLASTASDNDNF 421
QY      174 LSGERQR--KRH-----KSDSISLSFDESIALCVIREICCRSSSSSEST 215
DB      422 VTPEPRTRRHPTNQARASKKKPKVVFSSDES-----SEEDLSAEMT 464

RESULT 10
US-09-833-245-1731
; Sequence 1731, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1731
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-833-245-1731

Query Match      8.5%; Score 93; DB 11; Length 485;
Best Local Similarity 21.9%; Pred. No. 1.9;
Matches 51; Conservative 39; Mismatches 77; Indels 66; Gaps 12;

QY      22 SEQETLVPRKP---LLLKLKSVGAQKDTYTMKEVLFYQYIMTKRLYDEKQCHIVYCSN 78
DB      259 SDPELGVBEPPHTIMKQLLSYITKDKQTESLVEKL--CQRFRTSRT-ERQQRDLAYCVS 315
QY      79 -----DILGLDFGVPS--FSVKEHKIYTMIRNVLVAVNQQESSDSGTSV 121
DB      316 QLPLTERGLRKLNDNFDFCGDKLSDSIFSA-----FLSVGKLRGAKPEGKAI 365
QY      122 SE-----NRCHLEG--GSDQKDLVQELQEKPS--SHLVSRPSTSSRRRAISETEENSDE 173
DB      366 IDEFEQKLKACHTRGLDGIKELEIGQAGSQAPSA---KKPSTGSRVQPLASTASDNDNF 421
QY      174 LSGERQR--KRH-----KSDSISLSFDESIALCVIREICCRSSSSSEST 215
DB      422 VTPEPRTRRHPTNQARASKKKPKVVFSSDES-----SEEDLSAEMT 464

RESULT 11
US-09-833-245-1732
; Sequence 1732, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
```

```

; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1732
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-1732

Query Match      8.5%; Score 93; DB 11; Length 485;
Best Local Similarity 21.9%; Pred. No. 1.9;
Matches 51; Conservative 39; Mismatches 77; Indels 66; Gaps 12;

Qy 22 SEQETLVRRKP---LLKLKLSVGAQKDTYTMKEVLFYQIMTKRLYDEKQOHIVYCSN 78
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 259 SDFELGVEEPPHTIMKQLLSYTKDKQTESLVEKL---CQRPRTSRT-ERQORDLAYCVS 315
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 79 -----DLGDLFGVPS-FSVVHEHRIYTMIRLVNVLVNVNQESSDSGTSV 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 316 QLEPLTERGLRKLMDNDFCFGKLSDESIFSA-----FLSVVGLRGAKPEGKAI 365
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 122 SE-----NRCHLEG--GSDOKDLVQELQEEKSPSSHLVSRPSTSSRRRAISETEENSDE 173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 366 IDFEQKLRACHTRGDLGDKLEIGQAGSQRPASA-----KKPSTGSRVQPLASTASDNDF 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 174 LSGERQR---KRH-----KSDSISLSFDESALCVIRETCRRSSSSSST 215
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 422 VTPEPRTRRRHNTQQRASKKKPKVVFSSDES-----SEEDLSAEMT 464

RESULT 12
US-08-945-038-6
; Sequence 6, Application US/08945038
; Publication No. US20020146423A1
; GENERAL INFORMATION:
; APPLICANT: Doidge, Christopher V.
; APPLICANT: Lee, Adrian
; APPLICANT: Radcliff, Fiona J.
; APPLICANT: Hocking, Diana M.
; APPLICANT: Webb, Elizabeth A.
; TITLE OF INVENTION: PROTECTIVE HELICOBACTER ANTIGENS
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/945,038
; FILING DATE: 23-DEC-1997
; PRIOR APPLICATION DATA: PCT/AU96/00225
; APPLICATION NUMBER: 19-APR-1996
; FILING DATE: 19-APR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN2575
; FILING DATE: 21-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN3931
; FILING DATE: 03-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PN7565
; FILING DATE: 16-JAN-1996

; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 017227/0133
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 433 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-945-038-6

Query Match      8.4%; Score 92.5; DB 8; Length 433;
Best Local Similarity 24.3%; Pred. No. 1.8;
Matches 36; Conservative 32; Mismatches 63; Indels 17; Gaps 7;

Qy 45 KDTYTMKEVLFYQIMTKRLYD--EKQOHIVYCSNDLLGDLFGVPSFSVKEHR--KIYT 100
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 286 KNNMIVAQAQKTYTINNRRVFCGLAKLDHVVF--KNNLYGMVFGNLNFDITSHKSCRLGK 344
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 101 MYRNLVVNVNQESSDSGTSVSENCHLEGSDQKDLVQELQEEKPSS-----HLVSRPS 156
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 WYVEG--AGKENFNTSGVRALESH--HASVHAENDLVKAVQEDHITDSKYLEHKVHLM 401
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 157 TSSRRRAISETEENSELSGERQKRHK 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 402 DSAK-----HVKENIDKMFYEQDELNK 424

RESULT 13
US-10-078-531-5
; Sequence 5, Application US/10078531
; Publication No. US20030049271A1
; GENERAL INFORMATION:
; APPLICANT: MARTIN, DENIS
; APPLICANT: RIOUX, STEPHANE
; APPLICANT: BRODEUR, BERNARD R.
; APPLICANT: HAMEL, JOSEE
; APPLICANT: RHEULT, PATRICK
; TITLE OF INVENTION: STREPTOCOCCUS PYOGENES POLYPEPTIDES AND CORRESPONDING
; FILE REFERENCE: PHARMA-18
; CURRENT APPLICATION NUMBER: US/10/078,531
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/269,840
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 5
; LENGTH: 963
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-10-078-531-5

Query Match      8.2%; Score 89.5; DB 14; Length 963;
Best Local Similarity 22.0%; Pred. No. 11;
Matches 46; Conservative 38; Mismatches 88; Indels 37; Gaps 10;

Qy 3 NTNMSVPTDGAVTTSQIPASQEITLVREKPIILLKLKLSVG--AQ----KDTYTMKEVLFY 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 743 NKAVSAYNNNSAIFKANVRKLEKEL-----DLTLGVEGKGLAQATVWQGVYLLKTPLEPL 797
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 57 LOYIMTKRLYDEKQOHIVYC--SNDLLG-----DLFGVPSFSVKEHRIYTM-----Y 103
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 798 PEYIGLNVYFDKSGKLIYALDMSDTIGCGKDAYGNPILNVNDEGYNHALAVATLADY 857
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 104 RNL---VVNVNQESS--DSGTSVSENCHLEGSGQKDLVQELQEEKSSSHLVERPSTSS 159
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 858 EGLDINKTILNLSQLTSIRQVPTAAVFRAG-----IFQAIQNAAEAEQLLPKPGTHS 911

```

ORGANISM: Streptococcus pyogenes
US-10-078-531-8
Query Match 8.1%; Score 89; DB 14; Length 969;
Best Local Similarity 22.6%; Pred. No. 12;
Matches 45; Conservative 35; Mismatches 87; Indels 32; Gaps 9;
QY 13 AVTTSQIPASBOETLVRPKPLLLKLSVG--AQ-----KDTYTMKEVLFYLYQYIMTKRLY 66
DB 754 AYNSAIKKANVKRLEKELDLLTGLVEGKGLAQTWVGYYLLKTPLPPEYVIGLNVY 813
QY 67 DEKQOHIVYC--SNDLLG---DLFGVPFSVSKHRKIYTM-----YRNL---VVVN 110
DB 814 FDKSGKLIYALDMSDTIGEGQKDAYGNPILNVDEDSYHALAVATLADYEGLDIKTILN 873
QY 111 QOESS-DSGTSVSENRCHEGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISETEE 169
DB 874 SKLSQLTSSIRQVPTAAYHRAQ-----IFQAIGNAAAEQLLPKPGTHSEKSSSESAN 927
QY 170 NSD---ELSGERQKRHKHS 185
DB 928 SKDRGLQSNPKTNRGHSA 946
Search completed: March 14, 2004, 21:51:21
Job time : 59.3517 secs

160 RRAISATEENSD---ELSGERQKRHKHS 185
912 EKSSSESANSKDRGLQSNPKTNRGHSA 940
RESULT 14
US-10-078-531-3
Sequence 3, Application US/10078531
Publication No. US20030049271A1
GENERAL INFORMATION:
APPLICANT: MARTIN, DENIS
APPLICANT: RIOUX, STEPHANE
APPLICANT: BRODEUR, BERNARD R.
APPLICANT: HAMEL, JOSEE
APPLICANT: RHEAULT, PATRICK
TITLE OF INVENTION: STREPTOCOCCUS PYOGENES POLYPEPTIDES AND CORRESPONDING
TITLE OF INVENTION: DNA FRAGMENTS
FILE REFERENCE: PHARMA-18
CURRENT APPLICATION NUMBER: US/10/078,531
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/269,840
PRIOR FILING DATE: 2001-02-21
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 951
TYPE: PRT
ORGANISM: Streptococcus pyogenes
US-10-078-531-3
Query Match 8.1%; Score 89; DB 14; Length 951;
Best Local Similarity 22.6%; Pred. No. 12;
Matches 45; Conservative 35; Mismatches 87; Indels 32; Gaps 9;
QY 13 AVTTSQIPASBOETLVRPKPLLLKLSVG--AQ-----KDTYTMKEVLFYLYQYIMTKRLY 66
DB 736 AYNSAIKKANVKRLEKELDLLTGLVEGKGLAQTWVGYYLLKTPLPPEYVIGLNVY 795
QY 67 DEKQOHIVYC--SNDLLG---DLFGVPFSVSKHRKIYTM-----YRNL---VVVN 110
DB 796 FDKSGKLIYALDMSDTIGEGQKDAYGNPILNVDEDSYHALAVATLADYEGLDIKTILN 855
QY 111 QOESS-DSGTSVSENRCHEGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISETEE 169
DB 856 SKLSQLTSSIRQVPTAAYHRAQ-----IFQAIGNAAAEQLLPKPGTHSEKSSSESAN 909
QY 170 NSD---ELSGERQKRHKHS 185
DB 910 SKDRGLQSNPKTNRGHSA 928

RESULT 15
US-10-078-531-8
Sequence 8, Application US/10078531
Publication No. US20030049271A1
GENERAL INFORMATION:
APPLICANT: MARTIN, DENIS
APPLICANT: RIOUX, STEPHANE
APPLICANT: BRODEUR, BERNARD R.
APPLICANT: HAMEL, JOSEE
APPLICANT: RHEAULT, PATRICK
TITLE OF INVENTION: STREPTOCOCCUS PYOGENES POLYPEPTIDES AND CORRESPONDING
TITLE OF INVENTION: DNA FRAGMENTS
FILE REFERENCE: PHARMA-18
CURRENT APPLICATION NUMBER: US/10/078,531
CURRENT FILING DATE: 2002-02-21
PRIOR APPLICATION NUMBER: 60/269,840
PRIOR FILING DATE: 2001-02-21
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 969
TYPE: PRT

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2004, 21:55:40 ; Search time 916.834 Seconds
(without alignments)
2410.251 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 398

Sequence: 1 ARAYXKXFMFMRKAGKWC.....ARLXPCGKVGMDVRRRWS 74

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US10057510/runat_09032004.162235.25991/app.query.fasta_1.654
-DB=BST -OPMT=fastap -SUFFIX=std.rst -MINMATCH=0.1 -LOPEXT=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10057510@cgn.1.1.2921@runat_09032004.162235.25991 -NCPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: em_eatba.*
- 2: em_esthum.*
- 3: em_estin.*
- 4: em_estmu.*
- 5: em_estov.*
- 6: em_estpl.*
- 7: em_estro.*
- 8: em_htc.*
- 9: gb_est1.*
- 10: gb_est2.*
- 11: gb_htc.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: em_estfun.*
- 16: em_estom.*
- 17: em_gss_hum.*
- 18: em_gss_inv.*
- 19: em_gss_pln.*
- 20: em_gss_vrt.*
- 21: em_gss_fun.*
- 22: em_gss_mam.*
- 23: em_gss_mus.*
- 24: em_gss_pro.*
- 25: em_gss_rtd.*
- 26: em_gss_phg.*
- 27: em_gss_vrl.*
- 28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	331	83.2	1036	14	CD246058	CD246058 AGENCOURT
2	177	44.5	440	14	N28611	N28611 YX38F03.r1
3	176	44.2	1201	9	AL556102	AL556102 AL556102
4	165	41.5	298	10	BF989784	BF989784 PM2-GN001
5	165	41.5	377	10	BE302779	BE302779 ba75f06.Y
6	165	41.5	546	12	BM128107	BM128107 if08g03.Y
7	165	41.5	571	12	BM128076	BM128076 if08c12.Y
8	165	41.5	612	12	BG751548	BG751548 602730647
9	165	41.5	736	12	BM041856	BM041856 603615904
10	165	41.5	830	12	BM042118	BM042118 603615742
11	165	41.5	839	12	BG333546	BG333546 602460016
12	165	41.5	866	14	CD105325	CD105325 AGENCOURT
13	165	41.5	869	10	BF026125	BF026125 60169811
14	165	41.5	872	13	BQ213692	BQ213692 AGENCOURT
15	165	41.5	892	13	BQ718261	BQ718261 AGENCOURT
16	165	41.5	897	13	BU541223	BU541223 AGENCOURT
17	165	41.5	909	13	BU196508	BU196508 AGENCOURT
18	165	41.5	925	13	BU189380	BU189380 AGENCOURT
19	165	41.5	937	10	BK383862	BK383862 601273186
20	165	41.5	950	13	BQ944730	BQ944730 AGENCOURT
21	165	41.5	970	12	BG285692	BG285692 602380734
22	165	41.5	1024	10	BF337564	BF337564 602035315
23	165	41.5	1047	13	BU190249	BU190249 AGENCOURT
24	165	41.5	1087	10	BF978294	BF978294 602148259
25	165	41.5	1201	9	AL572268	AL572268 AL572268
26	165	41.5	1201	9	AL578516	AL578516 AL578516
27	162	40.7	418	12	BM696213	BM696213 UI-B-CL1-
28	162	40.7	486	12	BM689159	BM689159 UI-B-CQ1-
29	162	40.7	1057	12	BM562999	BM562999 AGENCOURT
30	162	40.7	1201	9	AL551959	AL551959 AL551959
31	162	40.7	1201	9	AL575410	AL575410 AL575410
32	162	40.7	1201	9	AL575424	AL575424 AL575424
33	158.5	39.8	827	12	BI552813	BI552813 603193789
34	158.5	39.8	948	13	BU156216	BU156216 AGENCOURT
35	157	39.4	678	10	AW577721	AW577721 RCO-BT056
36	156	39.2	569	10	BE073803	BE073803 RCO-BT056
37	156	39.2	1201	9	AL551978	AL551978 AL551978
38	155	38.9	716	13	BU633129	BU633129 UI-H-F11-
39	155	38.9	716	14	CD366140	CD366140 UI-H-F11-
40	154	38.7	336	10	AW750184	AW750184 RCO-BT056
41	154	38.7	497	10	BE073809	BE073809 RCO-BT056
42	154	38.7	629	10	AW602452	AW602452 RCO-BT056
43	154	38.7	630	10	AW602456	AW602456 RCO-BT056
44	154	38.7	653	10	AW577725	AW577725 RCO-BT056
45	154	38.7	658	10	AW602460	AW602460 RCO-BT056

ALIGNMENTS

RESULT 1
CD246058
LOCUS
DEFINITION
AGENCOURT_14095502 NIH_MGC_181 Homo sapiens CDNA clone
IMAGE:30377017 5', mRNA sequence.
ACCSSION
CD246058
VERSION
CD246058.1 GI:31006522
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1036)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Prepared by: Invitrogen Corp
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: NDAM440 row: j column: 02
 High quality sequence start: 19
 High quality sequence stop: 312.

FEATURES

source
 1..1036
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30377017"
 /issue type="White Matter"
 /dev stage="Unknown"
 /lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
 /clone_lib="NIH_MGC_181"
 /note="Vector: PCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV (destroyed); Library is oligo-dr primed and directionally cloned [EcoRV site is destroyed upon cloning]. Average insert size 1.42 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."
 Alignment Scores:
 Pred. No.: 1,468-29 Length: 1036
 Score: 331.00 Matches: 65
 Percent Similarity: 90.28% Conservative: 0
 Best Local Similarity: 90.28% Mismatches: 4
 Query Match: 93.17% Indels: 3
 DB: 14 Gaps: 1

US-10-057-510-2 (1-74) x CD246058 (1-1036)

Qy 6 LysMetPheSerMetPheMetGluLysAlaGlyLysCysTrpCysValMetProLysLeu 25
 |||||
Db 181 AAATGTTTTCATGTTTATGAAGAAGCTGGGAAGTCTGGTGTGAATGCCAAGCTC 240
 |||||
Qy 26 IleIleAspThrProPheSerIleValAlaProAlaLeuThrAlaValLeuSerCysGln 45
 |||||
Db 241 ATAATAGATACCTCTTCTCCATTGTTGCCCTGCTCAACTGCTGTTCTTTCTTGCCAG 300
 |||||
Qy 46 LeuArgCys-----SerLeuTrpLeuValGlyAlaArgLeu---ProCysGlyLysVa 63
 |||||
Db 301 NCNNNTTCGTTNNGTTTCCCTCTGGTGTGGGGGACGCGCTGTGCTCCATGTGGCAAGGT 360
 |||||
Qy 63 lGluGlyMetAspValTrpArgArgTrpSer 74
 |||||
Db 361 GGAGGCAATGGACGCTGTGGAGGAGCGCTGGAGC 394
 |||||

RESULT 2
 N28611
 LOCUS N28611
 DEFINITION YK38f03.r1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
 IMAGE:264029 5', mRNA sequence.

ACCESSION N28611.1 GI:1146947
VERSION N28611.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 440)
REFERENCE Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
 Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,

TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 285 1800
 Fax: 314 285 1810
 Email: est@watson.wustl.edu
 High quality sequence stops: 194
 Source: IMAGE Consortium, LLNL
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 900 Std Error: 0.00
 Seq primer: T7
 High quality sequence stop: 194.

FEATURES

source
 1..440
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:3873671"
 /db_xref="taxon:9606"
 /clone="IMAGE:264029"
 /sex="Male"
 /tissue type="melanocyte"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares melanocyte 2NbHM"
 /note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGGCCAGTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."

ORIGIN

Alignment Scores:
 Pred. No.: 2,27e-11 Length: 440
 Score: 177.00 Matches: 33
 Percent Similarity: 83.72% Conservative: 3
 Best Local Similarity: 76.74% Mismatches: 7
 Query Match: 44.47% Indels: 0
 DB: 14 Gaps: 0
 US-10-057-510-2 (1-74) x N28611 (1-440)
Qy 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 |||||
Db 87 TCCTTCATGCTACCCACCACTCAGTCGTGAGTCAGGCGAGCTTCGTGTTCCTCTGG 146
 |||||
Qy 52 LeuValGlyAlaArgLeu---ProCysGlyLysValGluGlyMetAspValTrpArgArg 71
 |||||
Db 147 CTGTGGGGGACGGCTGTTTCATGTGGCAAGGTGGAGGCACTGACCTGTGGAGGAGG 206
 |||||
Qy 72 ArgTrpSer 74
 |||||
Db 207 CGCTGGAGC 215
 |||||

RESULT 3
 AL556102
 LOCUS AL556102
 DEFINITION AL556102 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 cDNA clone CSDBK011YL07 5-PRIME, mRNA sequence.
ACCESSION AL556102
VERSION AL556102.2
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
 1201 bp mRNA linear EST 31-MAY-2003
 AL556102 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 cDNA clone CSDBK011YL07 5-PRIME, mRNA sequence.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1201)
 Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:12898460.

REFERENCE AUTHORS TITLE JOURNAL COMMENT

Contact: Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 628.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0K011CF04Q1&cluster=628.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CS0K011CF04Q1.

FEATURES source

1..1201
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0K011YLO7"
 /cell_type="HELA CELLS COT 25-NORMALIZED"
 /cell_line="HELA"
 /clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
 Pred. No.: 1,378-10 Length: 1201
 Score: 176.00 Matches: 32
 Percent Similarity: 81.40% Conservative: 3
 Best Local Similarity: 74.42% Mismatches: 8
 Query Match: 44.22% Indels: 0
 DB: 9 Gaps: 0

US-10-057-510-2 (1-74) x AL556102 (1-1201)

Qy 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 Db 867 TCCTTCATGCTACCCACCACTCAGTGTGAGTCAAGGACGCTGCTGTTCCCTCTGG 926
 Qy 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArg 71
 Db 927 CTGTGGGGGCACGGCTGTCTCATGTGCAAGGTGGAAGGCATGACGTGTGGAGGAG 986
 Qy 72 ArgTrpSer 74
 Db 987 CGCTGGAGC 995

RESULT 4

BF989784/c 298 bp mRNA linear EST 23-JAN-2001
 LOCUS PM2-GN0015-231000-005-cl2 GN0015 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BF989784
 VERSION BF989784.1 GI:12396109
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 298)
 Dias Neto, S., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brundstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F.F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

TITLE JOURNAL MEDLINE PUBMED COMMENT

Simpson, A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 10737800
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PM2&t2=PM2-GN0015-
 231000-005-cl2&t3=2000-10-23&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 298.

FEATURES source

1..298
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="GN0015"
 /note="Organ: placenta normal; Vector: puc18; Site 1:
 SmaI; Site 2: SmaI; A mini-library was made by cloning
 products derived from ORFESTES PCR (U.S. Letters Patent
 application No. 196,716 - Ludwig Institute for Cancer
 Research) profiles into the pUC 18 vector. Reverse
 transcription of tissue mRNA and cDNA amplification were
 performed under low stringency conditions."

ORIGIN

Alignment Scores:
 Pred. No.: 3,648-10 Length: 298
 Score: 165.00 Matches: 32
 Percent Similarity: 79.55% Conservative: 3
 Best Local Similarity: 72.73% Mismatches: 8
 Query Match: 41.46% Indels: 1
 DB: 10 Gaps: 0

US-10-057-510-2 (1-74) x BF989784 (1-298)

Qy 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 Db 198 TCCTTCATGCTACCCACCACTCAGTGTGAGTCAAGGACGCTGTTCCCTCTGG 139
 Qy 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArg 71
 Db 138 CTGTGGGGGCACGGCTGTCTCATGTGCAAGGTGGAAGGCATGACGTGTGGAGGAG 79
 Qy 71 ArgTrpSer 74
 Db 78 CGCTGGAGC 69

RESULT 5

BF989784/c 377 bp mRNA linear EST 14-JUL-2000
 LOCUS ba75f06.y1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:2906243 5',
 DEFINITION mRNA sequence.
 ACCESSION BF989784
 VERSION BF989784.1 GI:9186527
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 377)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE

JOURNAL
COMMENT

Unpublished (1999)
Other ESTs: ba75f06.x1
Contact: Robert Strausberg, Ph.D.
Email: cgabs-rc@mail.nih.gov
Tissue Procurement: ATCC/DCTP/DRP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40KP from Gibco
High quality sequence stop: 369.

FEATURES

source

1. 377
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2906243"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 20"
notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Alignment Scores:
Pred. No.: 5,19e-10 Length: 377
Score: 165.00 Matches: 32
Percent Similarity: 79.55% Conservative: 3
Best Local Similarity: 72.73% Mismatches: 8
Query Match: 41.46% Indels: 1
DB: 10 Gaps: 0

US-10-057-510-2 (1-74) x BE302779 (1-377)

QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTTP 51
Db 128 TCCTTCATGCTACCCACCTCAGTCTGAGGTCAGGCGAGCTTCGTTCCCTCTGG 187
QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 188 CTTGTGGGGGACGGCTGTCTCCATGTGGCAAGGTGGAAGGCATGGACGTGTGGAGGAG 247
QY 71 gAgTTPSer 74
Db 248 GCCTGGAGC 257

RESULT 6

BM128107

LOCUS

DEFINITION BM128107 546 bp mRNA linear EST 12-MAR-2002
if09a03.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:5676077 5' similar to TR:Q9Y5J5 Q9Y5J5 TDAG51/IPL
HOMOLOGUE 1. ; mRNA sequence.

ACCESSION

VERSION

BM128107

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemishka,I., Seearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Page,D., Wylie,T., Martin,J., Blinston,A.,
Schmitt,A., Theising,B., Ritter,B., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarishevli,R.,

TITLE
JOURNAL
COMMENT

Williams,T., Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: if09g03.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biochem.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 429.

FEATURES

source

1. 546
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5676077"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
notes="Organ: Pancreas; Vector: pSPORT1; Site 1: Not 1;
Site 2: Sal 1; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

ORIGIN

Alignment Scores:
Pred. No.: 9.08e-10 Length: 546
Score: 165.00 Matches: 32
Percent Similarity: 79.55% Conservative: 3
Best Local Similarity: 72.73% Mismatches: 8
Query Match: 41.46% Indels: 1
DB: 12 Gaps: 0

US-10-057-510-2 (1-74) x BM128107 (1-546)

QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTTP 51
Db 312 TCCTTCATGCTACCCACCTCAGTCTGAGGTCAGGCGAGCTTCGTTCCCTCTGG 371
QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 372 CTTGTGGGGGACGGCTGTCTCCATGTGGCAAGGTGGAAGGCATGGACGTGTGGAGGAG 431
QY 71 gAgTTPSer 74
Db 432 GCCTGGAGC 441

RESULT 7

BM128076

LOCUS

DEFINITION

BM128076

HOMOLOGUE 1. ; mRNA sequence.

[illegible]

SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: NDAM428 row: h column: 17
 High quality sequence stop: 618.

FEATURES
 source
 1..866
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30367384"
 /issue_type="Pituitary"
 /lab_host="DH10B-ron A (T1 and T5 phase resistances) "
 /clone_lib="NIH_MGC_179"
 /note="Organ: Brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2: NotI; Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.1 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.82e-09 Length: 866
 Score: 165.00 Matches: 32
 Percent Similarity: 79.55% Conservative: 3
 Best Local Similarity: 72.73% Mismatches: 8
 Query Match: 41.46% Indels: 1
 DB: 14 Gaps: 0
 US-10-057-510-2 (1-74) x CD105325 (1-866)
 QY 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 Db 507 TCCTTCATGCTACCCACCTCAGTGTGAGGTCAAGGCAGCTTCGTTGTTCCCTCTGG 566
 QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
 Db 567 CTTGTGGGGGACGGCTGTCTCCATGTGCAAGGTGGAAGCATGGACGTGTGGAGGAG 626
 QY 71 gArgTrpSer 74
 Db 627 GCGCTGGAGC 636

RESULT 13
LOCUS BF026125
DEFINITION 601669811F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3952874 5', mRNA sequence.
ACCESSION BF026125
VERSION BF026125.1 GI:10733837
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 869)
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory

SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC/DCTP/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 DNA Sequencing by: Invitrogen Corp
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Plate: LLCMI290 row: o column: 03
 High quality sequence stop: 703.

FEATURES
 source
 1..839
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4577114"
 /issue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant) "
 /clone_lib="NIH_MGC_20"
 /note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.74e-09 Length: 839
 Score: 165.00 Matches: 32
 Percent Similarity: 79.55% Conservative: 3
 Best Local Similarity: 72.73% Mismatches: 8
 Query Match: 41.46% Indels: 1
 DB: 12 Gaps: 0
 US-10-057-510-2 (1-74) x BG333545 (1-839)
 QY 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 Db 272 TCCTTCATGCTACCCACCTCAGTGTGAGGTCAAGGCAGCTTCGTTGTTCCCTCTGG 331
 QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
 Db 332 CTTGTGGGGGACGGCTGTCTCCATGTGCAAGGTGGAAGCATGGACGTGTGGAGGAG 391
 QY 71 gArgTrpSer 74
 Db 392 GCGCTGGAGC 401

RESULT 12
LOCUS CD105325
DEFINITION AGENCOURT_14021852 NIH_MGC_179 Homo sapiens cDNA clone IMAGE:30367384 5', mRNA sequence.
ACCESSION CD105325
VERSION CD105325.1 GI:30758499
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 866)

cDNA Library Arrayed by: The I.M.A.G.E.E. Consortium (LNLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E.E. Consortium/LNLNL at: image.llnl.gov
Plate: LNCM823 row: e column: 03
High quality sequence stop: 764

FEATURES

Source

```

High quality sequence, stop: 614.
Location/Qualifiers
1. 872
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6063433"
/tissue_type="melanotic"
/lab_host="PH10B (phage
/notes="Organ: skin; Vec
Site 2: SalI; cloned un
Average insert size 2 k
Technologies."

```

ORIGIN

ECORI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCACAG(G). Size-selected >500bp for agarose insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). "

Alignment Score:

US-10-057-S10-2 (1-74)	US-10-057-S10-2 (1-869)
BP026125	
DB:	10
Query Match:	41.46%
Best Local Similarity:	72.73%
Percent Similarity:	79.55%
Score, No.:	165/00
Alignment Scores:	
Length:	869
Matches:	32
Conservative:	3
Mismatches:	8
Indels:	1
Gaps:	0

	QY	31	TCTTCAATGCTACCCACCACCATCGTCAGGTCAAAGCAGCTTCGTTGTTCCCTCTGG	171
	DY	112	TCCTTCAATGCTACCCACCACCATCGTCAGGTCAAAGCAGCTTCGTTGTTCCCTCTGG	171
	QY	32	LeuValAlaProAlaLeuThrAlaValIleUserCysGlnLeuArgCysSerLeuTrp	51
	DY	112	LeuValAlaProAlaLeuThrAlaValIleUserCysGlnLeuArgCysSerLeuTrp	51
	QY	52	LeuValGlyAlaArgLeu-***ProCysGlyLyseValGluCluWeAspValTrpArg	71
	DY	172	CrrTGCGGGGCACGGCTGTGCTCCATGTGCAAGTGGAAAGGCATGAAGCTGTGGAGGAG	231

RESULT 15

Accession ID	BQ718261	BQ718261	BQ718261
LOCUS	BQ718261	892 bp	mRNA
DEFINITION	AGNCOURT 8294440	Luski_sympathetic trunk	Homo sapiens cDNA clone
	IMAGE:6194355	5', mRNA sequence.	
			linear
			EST 16-JUL-2002

FEATURES

source

BQ213692	872 bp	mRNA	linear	EST 02-MAY-2002									
LOCUS	AGNCOURT_7589137	NIH_MGC_72	Homo sapiens	cDNA clone IMAGE:6063433									
DEFINITION	5', mRNA sequence.												
ACCESSION	BQ213692												
VERSION	BQ213692.1	GI:20395092											
KEYWORDS	EST.												
SOURCE	Homo sapiens												
ORGANISM	Homo sapiens (human)												
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;												
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.												
	1 (bases 1 to 872)												
	NIH-MGC http://mgc.nci.nih.gov/ .												
AUTHORS	National Institutes of Health, Mammalian Gene Collection (MGC)												
TITLE	Unpublished (1999)												
COMMENT	Contact: Robert Strausberg, Ph.D.												
	Email: cgapbs-re@mail.nih.gov												
	Tissue procurement: ATCC/DCTD/DTF												
	cDNA Library Preparation: Life Technologies, Inc.												
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)												
	DNA Sequencing by: Agencourt Bioscience Corporation												
	Clone distribution by: Agencourt Bioscience Corporation												
	found through the I.M.A.G.E. Consortium information can be												
http://image.llnl.gov													
Plate: LLAM1336 Row: i Column: 02													

```

/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/clone_lib="Lupski sympathetic trunk"
/notes="vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3' and
5'-CACTAGTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

```

ORIGIN

Alignment Scores:

Pred. No.:	1.91e-09	Length:	892
Score:	165.00	Matches:	32
Percent Similarity:	79.55%	Conservative:	3
Best Local Similarity:	72.73%	Mismatches:	8
Query Match:	41.46%	Indels:	1
DB:	13	Gaps:	0

US-10-057-510-2 (1-74) x BQ718261 (1-892)

QY	32	SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp	51
DB	319	TCCTTCATGCTACCCACCTCAGTGTCTGAGGTCAAGGCAGCTTCGTTGTCCTCTGG	378
QY	52	LeuValGlyAlaArgLeu-***ProCysGlyIysValGluGlyMetAspValTrpArgAr	71
DB	379	CTTGTGGGGCAGCGCTGTGCTCCATGTGGCAAGGTGGAGGCATGGACGTGTGGAGGAG	438
QY	71	GArgTrpSer	74
DB	439	GCGTGGAGC	448

Search completed: March 15, 2004, 00:19:09
Job time : 921.834 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:33:39 ; Search time 25.2621 Seconds
(without alignments)
924.245 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 396

Sequence: 1 ARAYXKMFMEKAGKWC.....ARLXPGKVGMDVRRRWS 74

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mbc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_tvirus.*
- 16: sp_bacteriap.*
- 17: sp_archheap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	73	18.3	674	11 Q9DBJ2	Q9dbj2 mus musculus
2	69.5	17.5	411	13 Q7ZMU0	Q7zmu0 xenopus lae
3	69	17.3	1032	3 O14402	O14402 trichoderma
4	68	17.1	335	11 Q8R0S5	Q8r0s5 mus musculus
5	67	16.8	297	15 Q92U54	Q92u54 rhizobium m
6	64.5	16.2	308	16 Q88J70	Q88j70 pseudomonas
7	64.5	16.2	493	4 Q9NVH9	Q9nvh9 homo sapien
8	63	15.8	281	16 Q987W2	Q987w2 rhizobium l
9	62.5	15.7	636	3 O74560	O74560 schizosacch
10	62	15.6	100	11 Q9D7Q7	Q9d7q7 mus musculus
11	62	15.6	436	5 Q9N638	Q9n638 caenorhabdi
12	61.5	15.5	330	4 Q9H8I7	Q9h8i7 homo sapien
13	61.5	15.5	602	11 Q80VR0	Q80vr0 mus musculus
14	61.5	15.5	647	4 Q9IY20	Q9iy20 homo sapien
15	61.5	15.5	757	11 Q9D4F1	Q9d4f1 mus musculus
16	61.5	15.5	760	4 Q9NHV6	Q9nhv6 homo sapien

17	61	15.3	421	16 Q9X0D5	Q9x0d5 thermotoga
18	61	15.3	608	5 O62325	O62325 caenorhabdi
19	60.5	15.2	309	16 Q7U529	Q7u529 synchococc
20	60.5	15.2	321	16 Q7URJ7	Q7urj7 rhodopirell
21	60.5	15.2	406	16 Q8KF88	Q8kf88 chlorobium
22	60.5	15.2	439	2 Q9JQ61	Q9jq61 klebsiella
23	60.5	15.2	757	11 Q7TN61	Q7tn61 mus musculu
24	60	15.1	477	5 Q9V395	Q9v395 drosophila
25	60	15.1	553	5 Q9J367	Q9j367 caenorhabdi
26	60	15.1	598	12 Q8QRT6	Q8qrt6 chimpanzee
27	59.5	14.9	855	16 Q8FNR9	Q8fnr9 corynebacte
28	59.5	14.9	1211	3 Q9V751	Q9v751 pichia past
29	59	14.8	179	16 Q8YTG3	Q8ytc3 anabaena sp
30	59	14.8	320	5 O45814	O45814 caenorhabdi
31	59	14.8	392	17 Q8U0W1	Q8u0w1 pyrococcus
32	59	14.8	1135	12 Q9IKB5	Q9ikb5 tomato spot
33	59	14.8	1135	12 Q55647	Q55647 tomato spot
34	58.5	14.7	245	16 Q83N79	Q83n79 tropheryma
35	58.5	14.7	255	4 Q8WVY2	Q8wvy2 homo sapien
36	58.5	14.7	333	16 Q8DKT3	Q8dkt3 synchococc
37	58.5	14.7	801	12 Q9DQ82	Q9dq82 caprine her
38	58.5	14.7	2076	5 O01485	O01485 caenorhabdi
39	58	14.6	240	10 Q8H8T2	Q8h8t2 oryza sativ
40	58	14.6	351	16 Q885K1	Q885k1 pseudomonas
41	58	14.6	354	2 Q8VVB2	Q8vvb2 pseudomonas
42	58	14.6	460	16 Q899B9	Q899b9 clostridium
43	58	14.6	697	3 Q96UC6	Q96uc6 neurospora
44	58	14.6	741	16 Q8YBC5	Q8ybc5 bruceella me
45	58	14.6	787	16 Q8FX17	Q8fx17 bruceella su

ALIGNMENTS

RESULT 1

Q9DBJ2	PRELIMINARY;	PRT;	674 AA.
ID	Q9DBJ2		
AC	Q9DBJ2		
DT	01-JUN-2001 (TREMBLrel. 17, Created)		
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)		
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)		
DE	1300006023Rik protein.		
GN	1300006023Rik.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPRAIN=C57BL/6J; TISSUE=Liver;		
RX	MEDLINE=21085660; PubMed=11217851;		
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,		
RA	Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,		
RA	Lyons P., Marchionni I., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,		
RA	Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,		
RA	Havashizaki Y.		
RT	"Functional annotation of a full-length mouse cDNA collection."		
RL	Nature 409:685-690(2001).		
DR	EMBL; AK004920; BAB23671.1; -		
DR	MGI; MGI:1919014; 1300006023Rik.		

```
DR InterPro; IPR008973; C2_CaLB.
SQ SEQUENCE 674 AA; 72442 MW; BA290F25B6258E83 CRC64;

Query Match 18.3%; Score 73; DB 11; Length 674;
Best Local Similarity 24.1%; Pred. No. 0.77;
Matches 19; Conservative 9; Mismatches 23; Indels 28; Gaps 3;

QY 16 GKQCKWKMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXPCGK-----VEG 65
DB 3 GUGGCR-----EPGATGPDLPQLSPGVHCAWVGVRCVGRPGPAPPMAGVGRG 55
QY 66 M-----DWMRRRW 73
DB 56 LLASLEAFKSFRENWQRAW 74

RESULT 2
Q72WUO PRELIMINARY; PRT; 411 AA.
AC Q72WUO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to abhydrolase domain containing 3.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OC NCBI_TaxID=8355;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Klein S., Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046710; AA46710.1; -.
DR GO; GO:0016787; F.hydrolase activity; IEA.
DR InterPro; IPR000073; A/B hydrolase.
DR InterPro; IPR000379; Ser. esters.
DR InterPro; IPR000952; UPF0017.
DR Pfam; PFO0561; abhydrolase; 1.
DR PROSITE; PS01133; UPF0017; 1.
DR Hydrolase.
KW Hydrolase.
SQ SEQUENCE 411 AA; 46488 MW; AFE3EB1BC6B89C12 CRC64;

Query Match 17.5%; Score 69.5; DB 13; Length 411;
Best Local Similarity 24.6%; Pred. No. 1.3;
Matches 14; Conservative 12; Mismatches 24; Indels 7; Gaps 1;

QY 1 ARAYXMFMSFMKAGKCKMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARL 57
DB 23 ASAYVWFYRSYI-----CKPRLVSDLPFLFLRYCPVKEKRPFTWVCGGRI 72

RESULT 3
O14402 PRELIMINARY; PRT; 1032 AA.
AC O14402;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Beta-1,3 exoglucanase precursor.
OS Trichoderma harzianum
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.
OC NCBI_TaxID=5544;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=T-Y;
RA MEDLINE=99132288; PubMed=9931476;
RX Cohen-Kupiec R., Broglie K., Friesen D., Broglie R., Chet I.;
RT "Molecular Cloning and Characterization of a novel Beta-1,3-
RT exoglucanase related to Mycoparasitism of Trichoderma harzianum.";
RL Gene 226:147-154(1999).
```

```
DR EMBL; AJ002397; CAA05375.1; -.
DR PIR; T43257; T43257.
DR InterPro; IPR002889; WSC.
DR Pfam; PF01822; WSC; 2.
DR SMART; SM00321; WSC; 2.
KW SIGNAL.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 1032 BETA-1,3 EXOGLUCANASE.
SQ SEQUENCE 1032 AA; 107912 MW; F3E7390B3C398EEA CRC64;

Query Match 17.3%; Score 69; DB 3; Length 1032;
Best Local Similarity 32.2%; Pred. No. 3.9;
Matches 19; Conservative 7; Mismatches 21; Indels 12; Gaps 3;

QY 12 MEKAGKCKWKMPKLIIDTPF-SIVAPALTAVLSCQLRCSLWLVGARLXPCGKVGMDVW 69
DB 864 VEYSGEWC-----DTKFGNGGSPASDGSACTWTCS----GAPQETCGGNRLDVI 911

RESULT 4
Q8R0S5 PRELIMINARY; PRT; 335 AA.
ID Q8R0S5
AC Q8R0S5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OC ABCG4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026477; AA26477.1; -.
DR MGD; MGI:1890594; Abcg4.
DR Hypothetical protein.
KW NON TER 1
FT NON TER 1
SQ SEQUENCE 335 AA; 37230 MW; CD3AD5FB0D3D9EDD CRC64;

Query Match 17.1%; Score 68; DB 11; Length 335;
Best Local Similarity 27.5%; Pred. No. 1.7;
Matches 19; Conservative 10; Mismatches 28; Indels 12; Gaps 3;

QY 9 SMFMKAGKCKWK-----MPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXPCGKV 63
DB 270 AVFMREHLNYYTLKAYVLAKTMADVDFQVVC-----VVYCSIVYWLWVSPSLPGGRA 325
QY 64 E---GMDVW 69
DB 326 QLLACAVW 334

RESULT 5
Q92U54 PRELIMINARY; PRT; 297 AA.
ID Q92U54
AC Q92U54;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative sugar uptake ABC transporter permease protein.
OS RB1285 OR SM21459.
OC Rhizobium meliloti (Sinorhizobium meliloti).
OC Plasmid pSymb (megaplasmid 2).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OC NCBI_TaxID=382;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
```



```

RX MEDLINE=21396508; PubMed=11481431;
RA Pinan T.M., Weidner S., Wong K., Buhmester J., Chain P., Guzy J.,
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Guzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).
DR EMBL; AL603646; CAC49685.1; -.
DR FIR; E96002; E96002.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp. 1.
DR PROSITE; PS00402; BPD TRANSP INN_MEMBER; 1.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 297 AA; 32956 MW; 7A8141D5BED21774 CRC64;

Query Match 16.8%; Score 67; DB 16; Length 297;
Best Local Similarity 23.5%; Pred. No. 2;
Matches 16; Conservative 12; Mismatches 40; Indels 0; Gaps 0;

Qy 7 MFSPMEKAGKCKWPKLIIDTPFSIVAPALTAVLSQRLSLWVGARLXPKQVGE 66
Db 143 LVALGLVEKGVANTADPSILGTIFIDVWVTPFNVLLIAGLQLEPIEYEAADVSGV 202

Qy 67 DVWRERWS 74
Db 203 POWKRFWS 210

RESULT 6
Q88JTO PRELIMINARY; PRT; 308 AA.
ID Q88JTO;
AC Q88JTO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hydrolase, alpha/beta fold family.
GN P22567.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Reinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA Utterback T., Rizzo M., Lee K., Kosack D., Moesli D., Wedler H.,
RA Lauber J., Stjepandic D., Hobeisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808 (2002).
DR EMBL; AS016783; AAN68176.1; -.
DR TIGR; P22567; -.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR InterPro; IPR000073; A/b hydrolase.
DR InterPro; IPR000379; Ser esterase.
DR Pfam; PF00561; abhydrolase; 1.
KW Complete proteome.
SQ SEQUENCE 308 AA; 34218 MW; 311D83E4A1876942 CRC64;

Query Match 16.2%; Score 64.5; DB 16; Length 308;
Best Local Similarity 34.1%; Pred. No. 4.4;
Matches 14; Conservative 8; Mismatches 18; Indels 1; Gaps 1;

```

```

Qy 16 GKWCMPKLIIDTPFSIVAPALTAVLSQRLSLWVGAR 56
Db 207 GYCWREDPLVLPPEPPLEREACDLIS-QIRCFLLYIFGR 246

RESULT 7
Q9NVH9 PRELIMINARY; PRT; 493 AA.
ID Q9NVH9;
AC Q9NVH9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein FLJ10724.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA Wagatsuma M., Hosoliri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Iehida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Negahari K., Masuho Y.,
RA Ninomiya K., Iwayanagi T.;
RT "NEDO human cDNA sequencing project."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK001586; BAA91772.1; -.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 57298 MW; B8C37B73A4FC46D4 CRC64;

Query Match 16.3%; Score 64.5; DB 4; Length 493;
Best Local Similarity 28.3%; Pred. No. 7.1;
Matches 17; Conservative 10; Mismatches 24; Indels 9; Gaps 3;

Qy 18 CWCMPKLIIDTPFSIVAPALTAVLSQ--QLRSL--WLVGARLXPKQVKVGMDVRRRW 73
Db 105 CWC-----VLGTGQVAMVLLHTTISFCVAQFRSLLTLWLCSTLLSTLRLLQGVVEVKRW 159

RESULT 8
Q987W2 PRELIMINARY; PRT; 281 AA.
ID Q987W2;
AC Q987W2;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein ml16882.
GN ML16882.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338 (2000).
DR EMBL; AP003010; BAB53088.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 281 AA; 30421 MW; 4372A94245D1D700 CRC64;

Query Match 15.8%; Score 63; DB 16; Length 281;

```

Best Local Similarity 21.6%; Pred. No. 6.4; Matches 19; Conservative 11; Mismatches 28; Indels 30; Gaps 3;

QY 15 AGKCKWPKLIITDTFSIVAPALTAVALSCQLRCSLWLV-----53

Db 11 AGFAAKPFNNIVDAFLALSPMGLLVITLSLSARVWLPGRFVTLTLDNDYVRRPQMG 70

QY 54 GARLXPCGKVEG-----MDVWRRRW 73

Db 71 GGLWLP-GRDAILKELTLEFWRRAW 97

RESULT 9

O74560 PRELIMINARY; PRT; 636 AA.

AC O74560;

DT 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)

DE Hypothetical zinc finger protein.

GN SPCC970.07C.

OS Schizosaccharomyces pombe (Fission Yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972h-;

RA Murphy L., Harris D., Wood V., Rajandream M.A., Barrell B.G.;

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL031530; CAA20700.1; -

DR PIR; T41673; T41673.

DR GeneDB SPombe; SPCC970.07c; -

SQ SEQUENCE 636 AA; 73292 MW; C3F8306DD764CFB CRC64;

Query Match 15.7%; Score 62.5; DB 3; Length 636;

Best Local Similarity 30.8%; Pred. No. 17; Matches 16; Conservative 8; Mismatches 19; Indels 9; Gaps 2;

QY 28 DTFSIVAPALTAVALSCQLRCS-----LWLVGARLXPCGKVGMDVWRR 71

Db 244 EQPFSISSPATQLGVVKETSLRPPLNDIWSGLRIVD-PNIESISLWKR 294

RESULT 10

O9D7Q7 PRELIMINARY; PRT; 100 AA.

AC O9D7Q7;

DT 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)

DE 2210420N1ORik protein.

GN 2210420N1ORIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Stomach;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Peole G., Quackenbush J.,

RA Schirml L.M., Staubli P., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AK008991; BAB26011.1; -

DR MGD; MGI:1917387; 2210420N1ORik.

SQ SEQUENCE 100 AA; 10425 MW; E7C4A71C8DE08DE3 CRC64;

Query Match 15.6%; Score 62; DB 11; Length 100;

Best Local Similarity 37.5%; Pred. No. 3; Matches 12; Conservative 4; Mismatches 12; Indels 4; Gaps 1;

QY 42 LSCQLRCSLWLVGARLXPCGKVGMDVWRRRW 73

Db 20 LACRALSGARAAAGPRPPCGK-----QLWRRRW 47

RESULT 11

Q9N638 PRELIMINARY; PRT; 436 AA.

ID Q9N638;

AC Q9N638;

DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE F37D6.6 protein.

GN F37D6.6.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RA Wilkinson J.;

RL Submitted (JUG-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "Genome sequence of the nematode C.elegans: A platform for

RT investigating biology.";

RL Science 282:2012-2018(1998).

RN [3]

RP SEQUENCE FROM N.A.

RA McMurray A.A.;

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.

DR EMBL; 279600; CAB70221.1; -

DR EMBL; 275540; CAB70221.1; JOINED.

DR EMBL; 275540; CAB70215.1; -

DR EMBL; 279600; CAB70215.1; JOINED.

DR PIR; B87906; B87906.

DR PIR; T22995; T22995.

DR WormPep; F37D6.6; CE24960.

DR GO; GO:0005622; C:intracellular; IEA.

DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.

DR InterPro; IPR001132; DwarfIn.

DR InterPro; IPR003619; DwarfIn.A.

DR InterPro; IPR008984; SMAD_FHA.

DR Pfam; PF03166; MH2; 1.

DR SMART; SM00523; DWA; 1.

DR SMART; SM00524; DWA; 1.

SQ SEQUENCE 436 AA; 49496 MW; E0C71263BC580EEE CRC64;

Query Match 15.6%; Score 62; DB 5; Length 436;

Best Local Similarity 27.0%; Pred. No. 13; Matches 20; Conservative 8; Mismatches 28; Indels 18; Gaps 3;

QY 14 KAGKWCWPKL-----IIDT-----PFSIVAPALTAVALSCQLRCSLW-----LVGA 55

```
Db 268 ESNRSWAQWTRYEKEQIGDVTWMLHGPPAAVGVLSKSVHDAQLECSPWDLKNVSPALIR 327
QY 56 RLXPCKGVEGMDVW 69
Db 328 QADPIGSTNPEDVW 341

RESULT 12
Q9H8I7 PRELIMINARY; PRT; 330 AA.
AC Q9H8I7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ13593.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Isegai T., Ota T., Hayaashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H., Sugawara M.,
RA Wagaesuma M., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuho Y.,
RA Nimomiya K., Iwananagi T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK023655; BAB14629.1; -.
KW Hypothetical protein.
SQ SEQUENCE 330 AA; 38958 MW; EBA21A1AC1455C2E CRC64;

Query Match 15.5%; Score 61.5; DB 4; Length 330;
Best Local Similarity 28.2%; Pred. No. 12;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;

QY 4 YXKQFSMFV--EKAGKWCCKMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXP-- 59
Db 134 YSKNLSLWMNFQPPSKAW-RASQMTFFFLFFPSFTGVL-CTLAITW----RLKPSA 187

QY 60 -CGKVEGMDVW 69
Db 188 DCGPFRGLPLF 198

RESULT 13
Q80VR0 PRELIMINARY; PRT; 602 AA.
AC Q80VR0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to hypothetical protein FLJ13593 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046390; AAH46390.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 602 AA; 70147 MW; FC677F138F2D0E1 CRC64;

Query Match 15.5%; Score 61.5; DB 11; Length 602;
Best Local Similarity 30.9%; Pred. No. 21;
Matches 21; Conservative 10; Mismatches 26; Indels 11; Gaps 5;

Db 268 ESNRSWAQWTRYEKEQIGDVTWMLHGPPAAVGVLSKSVHDAQLECSPWDLKNVSPALIR 327
QY 56 RLXPCKGVEGMDVW 69
Db 328 QADPIGSTNPEDVW 341

RESULT 14
Q8IY20 PRELIMINARY; PRT; 647 AA.
AC Q8IY20;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ13593.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038118; AAH38118.1; -.
KW Genew; HGNC:22999; TMC5.
KW Hypothetical protein.
SQ SEQUENCE 647 AA; 75455 MW; FBB38E4249172F41 CRC64;

Query Match 15.5%; Score 61.5; DB 4; Length 647;
Best Local Similarity 28.2%; Pred. No. 23;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;

QY 4 YXKQFSMFV--EKAGKWCCKMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXP-- 59
Db 451 YSKNLSLWMNFQPPSKAW-RASQMTFFFLFFPSFTGVL-CTLAITW----RLKPSA 504

QY 60 -CGKVEGMDVW 69
Db 505 DCGPFRGLPLF 515

RESULT 15
Q9D4F1 PRELIMINARY; PRT; 757 AA.
AC Q9D4F1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 432443L08Rik protein (Tmc5 protein).
GN 432443L08Rik OR TMC5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavali T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schirkl L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
```

RA	Custincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S., Hayashizaki Y.;
RT	"Functional annotation of a full-length mouse cDNA collection.";
RN	Nature 409:685-690(2001).
[2]	
RP	SEQUENCE FROM N.A.
RC	STRAIN=B6;
RR	Keresztes G., Mutai H., Heller S.;
RT	"TMC and EVER genes belong to a larger novel family, the TMC gene family encoding transmembrane proteins.";
RL	BMC Genomics 4:24-24(2003)
DR	EMBL; AK016573; BAB30314.1; --
DR	EMBL; AY263159; AAP78774.1; --
DR	MGI; 1921674; 4932443J08Rik.
SQ	SEQUENCE 757 AA; 87667 MW; 320D61A71CF51DAE CRC64;
	Query Match 15.5%; Score 61.5; DB 11; Length 757;
	Best Local Similarity 30.9%; Pred. No. 27;
	Matches 21; Conservative 10; Mismatches 26; Indels 11; Gaps 5
Qy	4 YXKWFSPFM--EKAGKCKWCPKLIIPTPSIVAPALTAVLSQLRCSILWGVARLXP-- 59
Ddb	561 YKVNVLNMNFPPSKAW-RASQMIFPIILLFPFSFTGLV-CTLAITIW----RLKPSA 614
Qy	60 -CGKVEGM 66
Dd	615 DCGPFRGL 622

Search completed: March 14, 2004, 21:47:43
Job time : 28.2621 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 19:12:02 ; Search time 7.4 Seconds
(without alignments)
520.702 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 398

Sequence: 1 ARAYXKMFSPMEKAGKWC.....ARLXPGCKVGMVDVRRRWS 74

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65.5	16.5	397	1 YKQ3 CAEEL	P34298 caenorhabdi
2	62.5	15.7	293	1 FXK6 HUMAN	Q9nrd1 homo sapien
3	58.5	14.7	224	1 FX30 HUMAN	Q9h4m3 homo sapien
4	58.5	14.7	1822	1 ITP4 HUMAN	P16144 homo sapien
5	58	14.6	214	1 CMB2 DIACA	Q42498 dianthus ca
6	57.5	14.4	339	1 ORUR PSEAE	P72171 pseudomonas
7	57	14.3	285	1 REP1 BUCBP	Q95b46 buchnera ap
8	57	14.3	2769	1 THYG BOVIN	P01267 bos taurus
9	56	14.1	321	1 SPPA ARATH	Q9sel7 arabidopsis
10	56	14.1	542	1 YLM2 SCHPO	Q9p378 schizosacch
11	56	14.1	553	1 GLP2 HUMAN	Q95838 homo sapien
12	55.5	13.9	85	1 YS22 PSEAE	Q51483 pseudomonas
13	55	13.8	267	1 PTPC ECOLI	P42910 escherichia
14	55	13.8	926	1 PERT PIG	P09933 sus scrofa
15	54.5	13.7	711	1 PVP2 SCHPO	P32586 schizosacch
16	54.5	13.7	1015	1 PPOL BOVIN	P18493 bos taurus
17	54	13.6	203	1 R33 GARRU	Q9aif9 carsonella
18	54	13.6	419	1 VS48 TERVC	P22048 tomato blac
19	54	13.6	465	1 SELA DESBA	P56372 desulfovibr
20	54	13.6	1388	1 CAIE HUMAN	P39059 homo sapien
21	53.5	13.4	572	1 ACSA BACSU	P39062 bacillus su
22	53	13.3	104	1 Y473 CHLTR	O84479 chlamydia t
23	53	13.3	297	1 NADC HUMAN	Q15274 homo sapien
24	53	13.3	299	1 NADC MOUSE	Q91x91 mus musculu
25	53	13.3	357	1 G6PT HUMAN	P35575 homo sapien
26	53	13.3	368	1 HIS8 AGRT5	Q8U9w3 agrobacteri
27	53	13.3	592	1 TAT2 YEAST	P38967 saccharomyc
28	52.5	13.2	287	1 YJWJ ECOLI	P39409 escherichia
29	52.5	13.2	424	1 VS48 TERVL	P22050 tomato blac
30	52	13.1	256	1 EX21 ARATH	Q9f181 arabidopsis
31	52	13.1	277	1 TNK5 HUMAN	P25942 homo sapien
32	52	13.1	366	1 HIS8 PSEST	Q9r100 pseudomonas
33	52	13.1	500	1 RIAB_CVMUD	P26627 murine coro

ALIGNMENTS

RESULT 1

YKQ3 CAEEL STANDARD; PRT; 397 AA.

AC P34298;
01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein C06E1.3 in chromosome III.
GN C06E1.3.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Bristol N2;

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Fraser A., Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Karshaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shownkeen R., Sims M., Smaldon N., Smith A., Smith M., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;

RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";

RL Nature 368:32-38(1994).

RN [2]

RP REVISIONS.

RA Waterston R.;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; L16559; AAA27934.2; -

DR WormPep; C06E1.3; CE30481.

DR Hypothetical protein; Transmembrane.

KW TRANSNEM 26 46 POTENTIAL.

FT TRANSNEM 57 77 POTENTIAL.

FT TRANSNEM 191 211 POTENTIAL.

FT TRANSNEM 218 238 POTENTIAL.

FT TRANSNEM 261 281 POTENTIAL.

SQ SEQUENCE 397 AA; 43801 MW; 3E6A1B2A8264A56 CRC64;

Query Match 16.5%; Score 65.5; DB 1; Length 397;

Best Local Similarity 30.2%; Pred. NO. 0.84;

Q8k9l3 buchnera ap
Q9ctg6 mus musculu
Q99n01 rattus norv
P34535 caenorhabdi
Q9nq11 homo sapien
P19751 m replicase
P31548 escherichia
Q92083 mus musculu
Q9qy81 mus musculu
P22051 tomato blac
Q9ln94 arabidopsis
P53047 saccharomyc

34 52 13.1 514 1 MVIN BUCAP
35 52 13.1 588 1 ATY1 MOUSE
36 52 13.1 723 1 S21C RAT
37 52 13.1 737 1 YNC2 CAEEL
38 52 13.1 1180 1 ATY1 HUMAN
39 52 13.1 7180 1 RIAB_CVMUD
40 51.5 12.9 232 1 THIQ ECOLI
41 51.5 12.9 239 1 CLDE MOUSE
42 51.5 12.9 364 1 WN16 MOUSE
43 51.5 12.9 424 1 VS48 TERVS
44 51 12.8 262 1 EXP7 ARATH
45 51 12.8 317 1 RTA1 YEAST

```
Matches 13; Conservative 8; Mismatches 19; Indels 3; Gaps 1;
QY 18 CWCMPKLIIDTFPSIVAPALTAVALSCQLRCSLWLVGARLXPC 60
    ||| ||| : : : : : : : : : : : : : : : : :
DB 199 CWCLSVVLMLFLPHNAYKIGILATGICLIACLVLL---LSPC 238

RESULT 2
FBX6_HUMAN STANDARD; PRT; 293 AA.
ID FBX6_HUMAN Q9UKT3;
AC Q9NRD1; Q9UKT3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE F-box only protein 6 (F-box/G-domain protein 2).
GN FBX6 OR FBX6 OR FBG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=200303060; PubMed=10531035;
RA Cenciarelli C., Chaur D.S., Guardavaccaro D., Parks W., Vidal M.,
RA Pagano M.;
RT "Identification of a family of human F-box proteins.";
RL Curr. Biol. 9:1177-1179(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20399565; PubMed=10945468;
RA Ilyin G.P., Rialland M., Pigeon C., Gugen-Guillouzo C.;
RT "cDNA cloning and expression analysis of new members of the mammalian
RL F-box protein family.";
RL Genomics 67:40-47(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshitaki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.J., Skalka U., Snailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex
CC (by similarity).
CC -!- SIMILARITY: Contains 1 F-box domain.
CC -!- SIMILARITY: Contains 1 FBA domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).
```

```
DR EMBL; AF129536; AAF04470.1; ALT_INIT.
DR EMBL; AF233223; AAF67153.1; -.
DR EMBL; BC020880; AAH20880.1; -.
DR Genew; HGNC:13585; FBX06.
DR MIN; 605647; -.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR007397; FBA.
DR InterPro; IPR008945; SKP1_Skp2.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PS0181; FBOX; 1.
KW Ub1 conjugation pathway.
FT DOMAIN 10 57
FT CONFLICT 18 18 E->D (IN REF. 1).
FT CONFLICT 51 51 M->L (IN REF. 1).
FT CONFLICT 61 61 E->K (IN REF. 1).
FT CONFLICT 96 96 E->N (IN REF. 1).
FT CONFLICT 114 114 E->D (IN REF. 1).
FT CONFLICT 123 123 D->E (IN REF. 1).
FT CONFLICT 132 132 Y->S (IN REF. 1).
FT CONFLICT 139 139 M->L (IN REF. 1).
FT CONFLICT 143 144 SQ->WE (IN REF. 1).
FT CONFLICT 149 152 VAEG->LADR (IN REF. 1).
FT CONFLICT 287 287 R->G (IN REF. 1).
SQ SEQUENCE 293 AA; 33932 MW; EA4235CD9CCD80FF CRC64;

Query Match 15.7%; Score 62.5; DB 1; Length 293;
Best Local Similarity 30.8%; Pred. No. 1.5;
Matches 16; Conservative 13; Mismatches 14; Indels 9; Gaps 2;
QY 21 KMPKLIIDTFPSIVAPALTAVALSCQLRCSLWLVGARLXPCGVGMVWRR 72
    : : : : : : : : : : : : : : : : : :
DB 15 ELPENILLELFTHV-PARQLLNCLVCSLWR-----DLIDLMTLWKK 57

RESULT 3
FBX30_HUMAN STANDARD; PRT; 224 AA.
ID FBX30_HUMAN Q9H4M3;
AC Q9H4M3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE F-box protein FBX30.
GN FBX30.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Paulson H.L., Koppenshafer S.L.;
RA "FBX30; a novel member of NFB42 class of Fbx genes.";
RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=B-cell;
RX MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Schaefer C.F., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Brownstein M.J., Usdin T.B., Toshitaki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalon D.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Bouffard G.G.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.J., Skalka U., Snailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation.
CC -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex
CC (by similarity).
CC -!- SIMILARITY: Contains 1 F-box domain.
CC -!- SIMILARITY: Contains 1 FBA domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@sib-sib.ch).
```

```
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Kzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Probably recognizes and binds to some phosphorylated
CC proteins and promotes their ubiquitination and degradation (By
CC similarity)
CC -!- SUBUNIT: Part of a SCF (SKP1-cullin-F-box) protein ligase complex
CC (By similarity).
CC -!- SIMILARITY: Contains 1 F-box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: AY007380; AAC09623.1; -
CC EMBL: BC007832; AA07832.1; -
CC InterPro: IPR001810; F-box.
CC InterPro: IPR007397; FBA.
CC Pfam: PF006645; F-box; 1.
CC Pfam: PF04300; FBA; 1.
CC SMART: SM00256; FBOX; 1.
CC PROSITE: PS0181; FBOX; 1.
CC UBL conjugation pathway.
CC DOMAIN 3 50 F-BOX.
CC SEQUENCE 224 AA; 25698 MW; 7016A8750CB6FB71 CRC64;
SQ
Query Match 14.7%; Score 58.5; DB 1; Length 224;
Best Local Similarity 28.8%; Pred. No. 3.8; Indels 9; Gaps 2;
Matches 15; Conservative 14; Mismatches 14;
Qy 21 KMPKLIIDTPSIVAPALTAVLSQRLCSLWLVGARLXPKGVGMDVWRRR 72
Db 8 ELPEVILLETHTV-PARQLLNCLVCLWR-----DLIDLVLTKRK 50
RESULT 4
ID ITB4 HUMAN STANDARD; PRT; 1822 AA.
AC P16144; O14690; O14691; O15339; O15340; O15341; Q9UIQ4;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Integrin beta-4 precursor (GP150) (CD104 antigen).
GN ITGB4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. (ISOFORM BETA-4A).
RX MEDLINE=90183973; PubMed=2311577;
RA Suzuki S., Naitoh Y.;
RT "Amino acid sequence of a novel integrin beta 4 subunit and primary
RT expression of the mRNA in epithelial cells."
RL EMBO J. 9:757-763(1990).
[2]
RP SEQUENCE FROM N.A. (ISOFORM BETA-4B).
RX MEDLINE=90183974; PubMed=2311578;
RA Hogervorst F., Kuikman I., von Dem Borne A.E.G.K., Sonnenberg A.;
RT "Cloning and sequence analysis of beta-4 cDNA: an integrin subunit
RT that contains a unique 118 kd cytoplasmic domain."
RL EMBO J. 9:765-770(1990).
[3]
RP SEQUENCE FROM N.A. (ISOFORM BETA-4C).
RP TISSUE=Pancreas;
RX
```

```
RX MEDLINE=91009492; PubMed=1976638;
RA Tamura R.N., Rozzo C., Starr L., Chambers J., Reichardt L.F.,
RA Cooper R.M., Quaranta V.;
RT "Epithelial integrin alpha 6 beta 4: complete primary structure of
RT alpha 6 and variant forms of beta 4."
RL J. Cell Biol. 111:1593-1604(1990).
[4]
RN SEQUENCE FROM N.A. (ISOFORMS BETA-4A; BETA-4B AND BETA-4C).
RX MEDLINE=9738298; PubMed=9194858;
RA Pulkkinen L., Kurtz K.S., Xu Y., Bruckner-Tuderman L., Uitto J.;
RT "Genomic organization of the integrin beta 4 gene (ITGB4): a
RT homozygous splice-site mutation in a patient with junctional
RT epidermolysis bullosa associated with pyloric atresia."
RL Lab. Invest. 76:823-833(1997).
[5]
RN SEQUENCE FROM N.A. (ISOFORMS BETA-4A; BETA-4B AND BETA-4C).
RX TISSUE=Lung;
RA MEDLINE=97311186; PubMed=9166594;
RA Iacovacci S., Gagnoux-Palacios L., Zambruno G., Meneguzzi G.,
RA D'Alessio M.;
RT "Genomic organization of the human integrin beta 4 gene."
RL Mamm. Genome 8:448-450(1997).
[6]
RN REVISIONS.
RA D'Alessio M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBSJ databases.
[7]
RN SEQUENCE FROM N.A. (ISOFORM BETA-4E).
RX MEDLINE=97350870; PubMed=9207246;
RA van Leusden M.R., Kuikman I., Sonnenberg A.;
RT "The unique cytoplasmic domain of the human integrin variant beta4E is
RT produced by partial retention of intronic sequences."
RL Biochem. Biophys. Res. Commun. 235:826-830(1997).
[8]
RN SEQUENCE OF 28-46.
RP MEDLINE=89251596; PubMed=2542022;
RA Kajiji S., Tamura R.N., Quaranta V.;
RT "A novel integrin (alpha E beta 4) from human epithelial cells
RT suggests a fourth family of integrin adhesion receptors."
RL EMBO J. 8:673-680(1989).
[9]
RP ALTERNATIVE SPLICING (ISOFORM BETA-4D).
RX MEDLINE=95073005; PubMed=7982032;
RA Clarke A.S., Lotz M.M., Mercurio A.M.;
RT "A novel structural variant of the human beta 4 integrin cDNA."
RL Cell Adhes. Commun. 2:1-6(1994).
[10]
RN VARIANTS EB-PA TYR-61; CYS-252; ARG-562 AND TRP-1281.
RX MEDLINE=99011257; PubMed=9792864;
RA Pulkkinen L., Rouan F., Bruckner-Tuderman L., Wallerstein R.,
RA Garzon M., Brown T., Smith L., Carter W.G., Uitto J.;
RT "Novel ITGB4 mutations in lethal and nonlethal variants of
RT epidermolysis bullosa with pyloric atresia: missense versus
RT nonsense."
RL Am. J. Hum. Genet. 63:1376-1387(1998).
[11]
RN VARIANT EB-PA GLY-245.
RP MEDLINE=98082996; PubMed=9422533;
RA Pulkkinen L., Kim D.U., Uitto J.;
RT "Epidermolysis bullosa with pyloric atresia: novel mutations in the
RT beta-4 integrin gene (ITGB4)."
RL Am. J. Pathol. 152:157-166(1998).
[12]
RN VARIANT EB-PA PRO-156.
RX MEDLINE=98206430; PubMed=9546354;
RA Pulkkinen L., Bruckner-Tuderman L., August C., Uitto J.;
RT "Compound heterozygosity for missense (L156P) and nonsense (R554X)
RT mutations in the beta-4 integrin gene (ITGB4) underlies mild,
RT nonlethal phenotype of epidermolysis bullosa with pyloric atresia."
RL Am. J. Pathol. 152:935-941(1998).
[13]
RN VARIANT EB-PA ARG-38.
RP MEDLINE=99111354; PubMed=9892956;
```



```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L40405; AAB05559.1; -
CC DR EMBL; L40805; AAB63903.1; -
CC DR PIR; T10715; T10715.
CC DR HSSP; F11746; 1NMN.
CC DR TRANSFAC; T03061; -
CC DR InterPro; IPR002487; TF_Kbox.
CC DR InterPro; IPR002100; TF_MADSbox.
CC DR Pfam; PF01486; K-box; 1.
CC DR Pfam; PF00319; SRF-TF; 1.
CC DR PRINTS; PR00404; MADSDOMAIN.
CC DR SMART; SM00432; MADS; 1.
CC DR PROSITE; PS00350; MADS_BOX_1; 1.
CC DR PROSITE; PS00666; MADS_BOX_2; 1.
CC KW Transcription regulation; DNA-binding; Nuclear protein.
CC FT DOMAIN 3 58
CC FT DOMAIN 93 165
CC FT K-BOX.
CC SQ SEQUENCE 214 AA; 24757 MW; E5B8A4013D55D374 CRC64;

Query Match 14.6%; Score 58; DB 1; Length 214;
Best Local Similarity 24.6%; Pred. No. 4.2;
Matches 14; Conservative 11; Mismatches 16; Indels 16; Gaps 2;

Qy 33 IVAPALTAVALSCQLRCSLWVGAR-----LXP-----CGKVEGMDVWRRW 73
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 28 IMKKAQLTVLCAKAVSLMSTTHLHLHVLSPGSLKQWDEYQKIEGVDLWRKQW 84

RESULT 6
ORUR_PSEAE
ID ORUR_PSEAE STANDARD; PRT; 339 AA.
AC P72171;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ornithine utilization regulator.
GN ORUR OR PA0831.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Hickey C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Stover C.K., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: Probably activates the argJ gene that encodes ornithine
CC acetyltransferase. Binds to its own promoter-operator region.
CC Probably binds ornithine.
CC -!- SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U73506; AAB94774.1; -
CC DR EMBL; AE004518; AAG04220.1; -
CC DR PIR; G83540; G83540.
CC DR InterPro; IPR000005; HTHARAC.
CC DR Pfam; PF00165; HTH_Arac; 2.
CC DR PRINTS; PR00032; HTHARAC.
CC DR SMART; SM00342; HTH_ARAC; 1.
CC DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; FALSE_NEG.
CC DR PROSITE; PS00124; HTH_ARAC_FAMILY_2; 1.
CC KW Transcription regulation; Activator; DNA-binding; Complete proteome.
CC FT DNA BIND 257 276
CC FT H-T-H MOTIF (BY SIMILARITY).
CC SQ SEQUENCE 339 AA; 37610 MW; F1376ED624E97270 CRC64;

Query Match 14.4%; Score 57.5; DB 1; Length 339;
Best Local Similarity 28.3%; Pred. No. 7.6;
Matches 15; Conservative 10; Mismatches 25; Indels 3; Gaps 2;

Qy 1 ARAYKXMFMEFME-KAGKCKWCKPKLIIDTPFSIVAPALTAVLSCQLRCSLW 52
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 180 ARYAIEVFGVAEFQPNHLSWPSEYLERPWALANPA--TVQMCEQCEALL 230

RESULT 7
REPL_BUCBP
ID REPL_BUCBP STANDARD; PRT; 285 AA.
AC Q89B46;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Probable replication-associated protein replA.
GN REPL OR BHP603.
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OG Plasmid pBep1.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135942;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=22426901; PubMed=12522265;
RA Van Ham R.C.H.J., Kamerbeek J., Palacios C., Rausell C., Abascal F.,
RA Bastolla U., Fernandez J.M., Jimenez L., Postigo M., Silva F.J.,
RA Tamames J., Viguera E., Iatorre A., Valencia A., Moran F., Moya A.;
RA "Reductive genome evolution in Buchnera aphidicola.";
RT Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -!- FUNCTION: This protein is essential for plasmid replication; it is
CC involved in copy control functions (By similarity).
CC -!- SIMILARITY: BELONGS TO THE INCFII REPA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF492591; AAO31542.1; -
CC DR InterPro; IPR003446; RepLtn.
CC DR Pfam; PF02387; IncFII_repA; 1.
CC KW Plasmid; DNA replication; Plasmid copy control; Complete proteome.
CC SQ SEQUENCE 285 AA; 33320 MW; 18C5B38854E862C CRC64;

Query Match 14.3%; Score 57; DB 1; Length 285;
Best Local Similarity 42.9%; Pred. No. 7.5;
Matches 15; Conservative 4; Mismatches 6; Indels 10; Gaps 3;

Qy 6 KMFSPMEFKAG--KC---W-----CMPKLIIDTP 30
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 RLISQFMEPMGTFCKKWKDKILGCMYPMKIMLTP 155

or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L40405; AAB05559.1; -
CC DR EMBL; L40805; AAB63903.1; -
CC DR PIR; T10715; T10715.
CC DR HSSP; F11746; 1NMN.
CC DR TRANSFAC; T03061; -
CC DR InterPro; IPR002487; TF_Kbox.
CC DR InterPro; IPR002100; TF_MADSbox.
CC DR Pfam; PF01486; K-box; 1.
CC DR Pfam; PF00319; SRF-TF; 1.
CC DR PRINTS; PR00404; MADSDOMAIN.
CC DR SMART; SM00432; MADS; 1.
CC DR PROSITE; PS00350; MADS_BOX_1; 1.
CC DR PROSITE; PS00666; MADS_BOX_2; 1.
CC KW Transcription regulation; DNA-binding; Nuclear protein.
CC FT DOMAIN 3 58
CC FT DOMAIN 93 165
CC FT K-BOX.
CC SQ SEQUENCE 214 AA; 24757 MW; E5B8A4013D55D374 CRC64;

Query Match 14.6%; Score 58; DB 1; Length 214;
Best Local Similarity 24.6%; Pred. No. 4.2;
Matches 14; Conservative 11; Mismatches 16; Indels 16; Gaps 2;

Qy 33 IVAPALTAVALSCQLRCSLWVGAR-----LXP-----CGKVEGMDVWRRW 73
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 28 IMKKAQLTVLCAKAVSLMSTTHLHLHVLSPGSLKQWDEYQKIEGVDLWRKQW 84

RESULT 6
ORUR_PSEAE
ID ORUR_PSEAE STANDARD; PRT; 339 AA.
AC P72171;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ornithine utilization regulator.
GN ORUR OR PA0831.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Hickey C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Stover C.K., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -!- FUNCTION: Probably activates the argJ gene that encodes ornithine
CC acetyltransferase. Binds to its own promoter-operator region.
CC Probably binds ornithine.
CC -!- SIMILARITY: BELONGS TO THE ARAC/XYLS FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```

DR DR InterPro; IPR000716; Thyroglobulin_1.
DR DR Pfam; PF00135; Coesterase; 1.
DR DR Pfam; PF00086; thyroglobulin_1; 10.
DR DR SMART; SM00211; TY; 10.
DR DR PROSITE; PS00484; THYROGLOBULIN_1; 9.
DR DR PROSITE; PS00941; CARBOXYLESTERASE_B_2; 1.
KW Glycoprotein; Repeat; Thyroid hormone; Thyroid hormones biosynthesis;
KW Iodination; Sulfation; Signal.
FT SIGNAL 1 19
FT CHAIN 20 2769
FT DOMAIN 31 92
FT DOMAIN 93 160
FT DOMAIN 161 297
FT DOMAIN 298 358
FT DOMAIN 604 657
FT DOMAIN 661 725
FT DOMAIN 726 921
FT DOMAIN 922 1073
FT DOMAIN 1074 1145
FT DOMAIN 1146 1210
FT DOMAIN 1458 1471
FT REPEAT 1472 1488
FT REPEAT 1489 1505
FT REPEAT 1513 1567
FT REPEAT 1605 1725
FT REPEAT 1726 1893
FT REPEAT 1894 1996
FT REPEAT 1997 2130
FT REPEAT 2131 2188
FT MOD_RES 24 24
FT MOD_RES 24 24
FT MOD_RES 2574 2574
FT MOD_RES 2588 2588
FT MOD_RES 2767 2767
FT MOD_RES 2767 2767
FT CARBOHYD 110 110
FT CARBOHYD 198 198
FT CARBOHYD 483 483
FT CARBOHYD 495 495
FT CARBOHYD 747 747
FT CARBOHYD 853 853
FT CARBOHYD 947 947
FT CARBOHYD 1140 1140
FT CARBOHYD 1365 1365
FT CARBOHYD 1776 1776
FT CARBOHYD 1870 1870
FT CARBOHYD 2014 2014
FT CARBOHYD 2123 2123
FT CARBOHYD 2251 2251
FT CARBOHYD 2296 2296
FT CONFLICT 1206 1206
FT CONFLICT 1206 1206
SQ SEQUENCE 2769 AA; 303218 MW; 1C7F227E9101DE2A CRC64;

Query Match 14.3%; Score 57; DB 1; Length 2769;
Best Local Similarity 27.0%; Pred. No. 69;
Matches 24; Conservative 6; Mismatches 25; Indels 34; Gaps 4;

Oy 16 GKWC-----KMPKLI-----DTFFSVAPALTAVL-----SC 44
Db 1186 GSCWCVLGSSEVFGTRVAGSQPACSPQCLPFFSVADVAGGAILCERASGLGAAGQRC 1245
Oy 45 QLRCSLWLVGARLXPCGKVEGMDVWRRRW 73
Db 1246 QLRCSS---QGYRSAPPEPLCSVQRRRW 1271

RESULT 9
SPPA ARATH
ID SPPA ARATH STANDARD; PRT; 321 AA.
AC Q9SEL7; Q49507;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Protease sppa, chloroplast precursor (EC 3.4.21.-).
```


Db 23 ELPWGPAPWGTSPSLSPHRCSLWAPGRPFLTLVLVSIKQVTCGSLLEETTRKWA 77

RESULT 12

Y522_PSEAE STANDARD; PRT; 85 AA.

AC Q51483;

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein PA0522.

GN PA0522

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PA01161;

RC MEDLINE=94362287; PubMed=7769251;

RA Arai H., Igarashi Y., Kodama T.;

RA "Structure and ANR-dependent transcription of the nir genes for denitrification from Pseudomonas aeruginosa.";

RT Biosci. Biotechnol. Biochem. 58:1286-1291(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;

RC MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P., Hickey M.J., Brinkman F.S.L., Hunsigle W.O., Kowalik D.J., Lagrou M., Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;

RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen.";

RT Nature 406:959-964(2000).

RL Nature 406:959-964(2000).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; D37883; BAA07125.1; -.

DR EMBL; AE004489; AAG03911.1; -.

DR PIR; JC2290.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 4 24 POTENTIAL.

FT TRANSMEM 27 47 POTENTIAL.

FT TRANSMEM 61 81 POTENTIAL.

SQ SEQUENCE 85 AA; 9097 MW; 18A1E43A65DCACOC CRC64;

Query Match 13.9%; Score 55.5; DB 1; Length 85;

Best Local Similarity 27.0%; Pred. No. 3.6;

Matches 17; Conservative 8; Mismatches 21; Indels 17; Gaps 4;

OY 18 CWCMPKLI-----DFFSIVAPALTAVLSQRLCSLWLVGALXPCGKVGMDV 68

DB 7 CWALLALAVTVGLVGGADSPW-----LLAAVLACAV-AGKWLIGRFMELAHAPAL-- 58

OY 69 WRR 71

DB 59 WRR 61

RESULT 13

PTPC_ECOLI STANDARD; PRT; 267 AA.

ID PTPC_ECOLI

AC P42910;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DE 28-FEB-2003 (Rel. 41, Last annotation update)

DE PTS system, N-acetylgalactosamine-specific IIC component 1 (EIIC-AGA)

DE (N-acetylgalactosamine-permease IIC component 1).

GN AGAC OR B3139.

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RC MEDLINE=9742617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12.";

RT Science 277:1453-1474(1997).

RL Science 277:1453-1474(1997).

RN [2]

RP DISCUSSION OF SEQUENCE.

RC MEDLINE=97086503; PubMed=8932697;

RA Reizer J., Rameisler T.M., Reizer A., Charbit A., Saier M.H. Jr.;

RA "Novel phosphotransferase genes revealed by bacterial genome sequencing: a gene cluster encoding a putative N-acetylgalactosamine metabolic pathway in Escherichia coli.";

RT Microbiology 142:231-250(1996).

CC -1- FUNCTION: This is a component of the phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS), a major carbohydrate active -transport system. The IICD domains contain the sugar binding site and the transmembrane channel; the IIA domain contains the primary phosphorylation site (the donor is phospho-Hpr); IIA transfers its phosphoryl group to the IIB domain which finally transfers it to the sugar.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential). Contains 1 PTS EIIC domain.

CC -1- SIMILARITY: Contains 1 PTS EIIC domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U18997; AAA57942.1; -.

DR EMBL; AE000395; AAC76173.1; -.

DR PIR; G65103; G65103.

DR EcGene; EGI2770; agac.

DR InterPro; IPR004700; Ptrans_sorb_IIC.

DR Pfam; PF03609; EII-Sor; 1.

DR TIGRFAMs; TIGR00822; EII-Sor; 1.

KW Phosphotransferase system; Sugar transport; Transmembrane; Inner membrane; Complete proteome.

FT TRANSMEM 11 31 POTENTIAL.

FT TRANSMEM 34 54 POTENTIAL.

FT TRANSMEM 67 87 POTENTIAL.

FT TRANSMEM 95 115 POTENTIAL.

FT TRANSMEM 142 162 POTENTIAL.

FT TRANSMEM 178 198 POTENTIAL.

FT TRANSMEM 211 231 POTENTIAL.

SQ SEQUENCE 267 AA; 28645 MW; C44F4D0827FE56C6 CRC64;

Query Match 13.8%; Score 55; DB 1; Length 267;

Best Local Similarity 24.1%; Pred. No. 13;

Matches 19; Conservative 10; Mismatches 22; Indels 28; Gaps 4;

OY 4 YXKMFNMEKAKGC-----WCKMPKLIIDTFFSIVAPALTAVLSQRLCSL 50

DB 114 FYSAPSLFMTKADKCAKEADTAAPSLNWTM--LIVASAYAVIA-----FLCTY 161

QY 51 WLVGARLXPCGKVEGMDVW 69
Db 162 LAQGAQ---ALVXAMPW 177

RESULT 14

```

PERT_PIG          STANDARD;          PRT;          926 AA.
AC  P09933;
DT  01-MAR-1989 (Rel. 10, Created)
DT  01-MAR-1989 (Rel. 10, last sequence update)
DT  15-MAR-2004 (Rel. 43, last annotation update)
DE  Thyroid peroxidase precursor (EC 1.11.1.8) (TPO).
GN  TPO.
OS  Sus scrofa (Pig).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX  NCBI_TaxID=9823;
RN  [1]
RN  SEQUENCE FROM N.A.
RP  MEDLINE=88007624; PubMed=3654642;
RX  Magnusson R.P., Gestautas J., Taurog A., Rapoport B.;
RA  "Molecular cloning of the structural gene for porcine thyroid
RT  peroxidase.";
RL  J. Biol. Chem. 262:13885-13888(1987).
RN  [2]
RN  SEQUENCE OF 595-926 FROM N.A.
RX  MEDLINE=87054611; PubMed=3780975;
RA  Magnusson R.P., Gestautas J., Seto P., Taurog A., Rapoport B.;
RT  "Isolation and characterization of a cDNA clone for porcine thyroid
RT  peroxidase.";
RL  FEBS Lett. 208:391-396(1986).
CC  -1- FUNCTION: Iodination and coupling of the hormonogenic tyrosines in
CC  thyroglobulin to yield the thyroid hormones T(3) and T(4).
CC  -1- CATALYTIC ACTIVITY: Iodide + H(2)O(2) = Iodine + 2 H(2)O.
CC  -1- COFACTOR: Binds 1 heme B (iron-protoporphyrin IX) group covalently
CC  and 1 calcium ion per heterodimer (By similarity).
CC  -1- PATHWAY: Thyroid hormone biosynthesis.
CC  -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC  -1- PTM: Heme is covalently bound through a H(2)O(2)-dependent
CC  autocatalytic process. Heme insertion is important for the
CC  delivery of protein at the cell surface (By similarity).
CC  -1- PTM: Cleaved in its N-terminal part (By similarity).
CC  -1- SIMILARITY: Belongs to the peroxidase family. XPO subfamily.
CC  -1- SIMILARITY: Contains 1 EGF-like domain.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; X04645; CAA28306.1; --
CC  PIR; A27416; OPGIT.
CC  HSP; P05164; 1CXP.
CC  InterPro; IPR002007; Anim_peroxidase.
CC  InterPro; IPR01881; EGF_CA.
CC  InterPro; IPR06209; EGF_like.
CC  InterPro; IPR06210; IEGF.
CC  InterPro; IPR002016; Peroxidase.
CC  InterPro; IPR000436; Sushi_SCR_CCP.
CC  Pfam; PF03098; An_peroxidase; I.
CC  Pfam; PF00008; EGF_1.
CC  Pfam; PF00084; sushi; 1.
CC  PRINTS; PR00457; ANPEROXIDASE.
CC  SMART; SM00032; CCP; 1.
CC  SMART; SM00181; EGF; 1.
CC  PROSITE; PS00022; EGF_1; FALSE_NEG.
CC  PROSITE; PS01186; EGF_2; FALSE_NEG.
CC  PROSITE; PS50026; EGF_3; 1.
CC  PROSITE; PS01187; EGF_CA; FALSE_NEG.

```

```

DR  PROSITE; PS00435; PEROXIDASE 1; 1.
DR  PROSITE; PS00436; PEROXIDASE 2; FALSE_NEG.
DR  PROSITE; PS0292; PEROXIDASE 3; 1.
KW  Oxidoreductase; Peroxidase; Iron; Heme; Transmembrane; Glycoprotein;
KW  EGF-like domain; Signal.
FT  SIGNAL 1 14
FT  CHAIN 15 926
FT  DOMAIN 19 844
FT  TRANSMEM 845 869
FT  POTENTIAL.
FT  CYTOPLASMIC (POTENTIAL).
FT  EGF-LIKE, CALCIUM-BINDING (POTENTIAL).
FT  DISTAL HISTIDINE (BY SIMILARITY).
FT  DISTAL ARGININE (BY SIMILARITY).
FT  HEME (COVALENT) (BY SIMILARITY).
FT  BINDING 238 238
FT  BINDING 238 238
FT  BINDING 398 398
FT  BINDING 240 240
FT  METAL 321 321
FT  METAL 321 321
FT  METAL 323 323
FT  METAL 323 323
FT  METAL 325 325
FT  METAL 327 327
FT  METAL 493 493
FT  DISULFID 142 158
FT  DISULFID 259 269
FT  DISULFID 263 286
FT  DISULFID 596 653
FT  DISULFID 694 719
FT  CARBOHYD 129 129
FT  CARBOHYD 277 277
FT  CARBOHYD 307 307
FT  CARBOHYD 342 342
SQ  SEQUENCE 926 AA; 100442 MW; 8549FF60F742C5E CRC64;
      13.8%; Score 55; DB 1; Length 926;
      Best Local Similarity 40.5%; Pred. No. 43;
      Matches 17; Conservative 3; Mismatches 14; Indels 8; Gaps 1;

QY 32 SIYALPALTAVLSQCLRSLLVVGAR-----LXPCGKVEG 65
      : : : : : : : : : : : : : : : : : : : :
Db 846 SVVSTALGALVLCGLAGLAWTVVCRWTHADARPLLPVGEGB 887
      : : : : : : : : : : : : : : : : : : : :

RESULT 15
ID  PYP2_SCHPO STANDARD;          PRT;          711 AA.
AC  P32586; Q9UR59; Q9UUG4;
DT  01-OCT-1993 (Rel. 27, Created)
DT  01-OCT-1993 (Rel. 27, Last sequence update)
DT  10-OCT-2003 (Rel. 42, Last annotation update)
DE  Protein-tyrosine phosphatase 2 (EC 3.1.3.48) (PTPase 2).
GN  PYP2 OR SPAC19D5.01.
OS  Schizosaccharomyces pombe (Fission yeast).
OC  Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC  Schizosaccharomycetales; Schizosaccharomycetaceae;
OC  Schizosaccharomycetes.
OX  NCBI_TaxID=4896;
RN  [1]
RN  SEQUENCE FROM N.A.
RX  MEDLINE=9309869; PubMed=1464319;
RA  Millar J.B.A., Russell P., Dixon J.E., Guan K.L.;
RT  "Negative regulation of mitosis by two functionally overlapping
RT  PTPases in fission yeast.";
RL  EMBO J. 11:4943-4952(1992).
RN  [2]
RN  SEQUENCE FROM N.A.
RX  MEDLINE=93078758; PubMed=1448087;
RA  Ohtsue S., Chernoff J., Hannig G., Hoffman C.S., Erikson R.L.;
RT  "The fission yeast genes pyp1+ and pyp2+ encode protein tyrosine
RT  phosphatases that negatively regulate mitosis.";
RL  Mol. Cell. Biol. 12:5571-5580(1992).
RN  [3]
RN  SEQUENCE FROM N.A.
RP  STRAIN=972;

```

RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayes J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gellies S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkhardt G., Aert R., Robben J., Grymonprez B.,
 RA Welljens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Foreburg S.L.,
 RA Cerruti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 [4]
 RP SEQUENCE OF 285-437 FROM N.A.
 RC STRAIN=968 h30;
 RX MEDLINE=2023868; PubMed=10759889;
 RA Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraquchi T.,
 RA Hiraoka Y.;
 RT "Large-scale screening of intracellular protein localization in living
 RT fission yeast cells by the use of a GFP-fusion genomic DNA library.";
 RL Genes Cells 5:169-190(2000).
 [5]
 RP SEQUENCE OF 485-519 FROM N.A.
 RX MEDLINE=93203106; PubMed=1668885;
 RA Yanagida M., Yamano H., Stone E.M., Kinoshita N., Yoshida T.,
 RA Shiozaki K.;
 RT "Protein phosphatases in cell division: how vital are they?";
 RL Princess Takamatsu Symp. 22:137-144(1991).
 [6]
 RP CHARACTERIZATION.
 RX MEDLINE=95385997; PubMed=7657164;
 RA Millar J.B.A., Buck V., Wilkinson M.G.;
 RA "Pyp1 and Pyp2 PRPases dephosphorylate an osmosensing MAP Kinase
 RT controlling cell size at division in fission yeast.";
 RL Genes Dev. 9:2117-2130(1995).
 CC -!- FUNCTION: Plays a role in inhibiting the onset of mitosis.
 CC Dephosphorylates sty1/spc1 and wsl1/spc2/sty2.
 CC -!- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to the protein-tyrosine phosphatase family.
 CC -!- SIMILARITY: Contains 1 rhodanese domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC EMBL; S51320; AAB24544.1; -;
 CC EMBL; X59599; CAA42167.1; -;
 CC EMBL; AB027789; BAA87093.1; -;
 CC EMBL; Z59531; CAB16711.1; -;
 CC FIR; S28391; S28391.
 CC PIR; T45160; T45160.

DR HSP; P18052; 1YFO.
 DR GenesDB Spombe; SPAC19D5.01; -;
 DR InterPro; IPR001763; Rhodanese-like.
 DR InterPro; IPR000387; TYR phosphatase.
 DR InterPro; IPR000422; TyP.
 DR Pfam; PF00581; Rhodanese; 1.
 DR Pfam; PF0102; Y-phosphatase; 1.
 DR PRINTS; PR00700; PRTYPHPTASE.
 DR SMART; SM00194; PTPC; 1.
 DR SMART; SM00450; RHOD; 1.
 DR PROSITE; PS0206; RHODANESE 3; 1.
 DR PROSITE; PS00383; TYR_PHOSPHATASE 1; 1.
 DR PROSITE; PS0056; TYR_PHOSPHATASE 2; 1.
 DR PROSITE; PS0055; TYR_PHOSPHATASE_PTP; 1.
 KW Cell division; Mitosis; Hydrolase.
 FT DOMAIN 21 130
 FT ACT_SITE 630 630
 FT PHOSPHOCYSTEINE INTERMEDIATE (BY
 FT SIMILARITY)
 FT IVGNA -> NRQC (IN REF. 2).
 FT I -> M (IN REF. 5).
 FT CONFLICT 487 487
 FT CONFLICT 515 515
 FT CONFLICT 519 519
 FT CONFLICT 636 636
 FT CONFLICT 668 670
 FT CONFLICT 675 676
 FT CONFLICT RS -> DT (IN REF. 2).
 SQ SEQUENCE 711 AA; 79356 MW; 15A12CDAE13D341E CRC64;
 Query Match 13.7%; Score 54.5; DB 1; Length 711;
 Best Local Similarity 41.9%; Pred. No. 38;
 Matches 13; Conservative 5; Mismatches 8; Indels 5; Gaps 1;
 QY 29 TPFSIVAPALTAIVLSQRLCSLWLVGARLXP 59
 DB 210 TPFSIHSPASSVRSCQS-----IYGSPLSP 235
 Search completed: March 14, 2004, 21:45:43
 Job time : 10.4 secs

A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Smb21459

A:Genome: plasmid

C:Superfamily: inner membrane protein ugpa

Query Match 16.8%; Score 67; DB 2; Length 297;

Best Local Similarity 23.5%; Pred. No. 1.4;

Matches 16; Conservative 12; Mismatches 40; Indels 0; Gaps 0;

Qy 7 MFSPFMEKAGKCKMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXPCGKVEGM 66

Db 143 LVALLGVEKGVANTADPSLILGTVFIDVWVTPFWLLLAGLQLIPEEIVEADYSGV 202

Qy 67 DVNRRRWS 74

Db 203 POWKRFWS 210

RESULT 3

A88533

hypothetical protein C06E1.3 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 02-Nov-2001

C:Accession: A88533

R:Anonymous; The C. elegans Sequencing Consortium.

Science 282, 202-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/c_elegans/ and www.sanger.ac.uk/Projects/C_ele

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: A88533

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-383 <STO>

A:Cross-references: GB:chr_III; PIDN:AAA27934.1; PID:g289628; GSPDB:GN00021

A>Note: putative

C:Genetics:

A:Map position: 3

Query Match 16.5%; Score 65.5; DB 2; Length 383;

Best Local Similarity 30.2%; Pred. No. 2.7;

Matches 13; Conservative 8; Mismatches 19; Indels 3; Gaps 1;

Qy 18 CWCCKPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXPC 60

Db 185 CWCLSVLMFLFPHNAYKSLATGSLIACLVLL--LSPC 224

RESULT 4

T41673

hypothetical zinc finger protein SPCC970.07c - fission yeast (*Schizosaccharomyces pombe*)

C:Species: *Schizosaccharomyces pombe*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000

C:Accession: T41673

R:Murphy, L.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, September 1998

A:Reference number: Z22009

A:Accession: T41673

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-636 <MUR>

A:Cross-references: EMBL:AL031530; PIDN:CAA20700.1; GSPDB:GN00068; SPDB:SPCC970.07c

A:Experimental source: strain 972h; cosmid c970

C:Genetics:

A:Gene: SPDB:SPCC970.07c

A:Map position: 3

C:Superfamily: Schizosaccharomyces hypothetical zinc finger protein SPCC970.07c

Query Match

Best Local Similarity

15.7%; Score 62.5; DB 2; Length 636;

Matches 10; Conservative 8; Mismatches 10; Indels 0; Gaps 0;

Matches 16; Conservative 8; Mismatches 19; Indels 9; Gaps 2;

Qy 28 DTPFSIVAPALTAVLSCQLRCS-----LWLVGARLXPCGKVEGMVRR 71

Db 244 EQPFSISSTATPQGVKRTSPLRPFLNDIWLGLRIVD-PNIESISLWKR 294

RESULT 5

T21920

hypothetical protein F37D6.7 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T21920

R:McMurray, A.

submitted to the EMBL Data Library, June 1996

A:Reference number: Z19487

A:Accession: T21920

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-249 <WIL>

A:Cross-references: EMBL:Z75540; PIDN:CAA99852.1; GSPDB:GN00019; CESP:F37D6.7

A:Experimental source: clone F37D6

C:Genetics:

A:Gene: CESP:F37D6.7

A:Map position: 1

A:Introns: 41/1; 62/1; 115/1; 176/1; 203/1

Query Match

Best Local Similarity

15.6%; Score 62; DB 2; Length 249;

Matches 20; Conservative 8; Mismatches 28; Indels 18; Gaps 3;

Qy 14 KAKCKCKMPKL-----IDT-----PFSIVAPALTAVLSCQLRCSLW-----LVGA 55

Db 172 ESNRSWAKMTRYERKEIGDVTWLGPFPAAGVLSKVHDAQLECSFMDLKNEVSPALIR 231

Qy 56 RLXPCGKVEGMV 69

Db 232 QADPIGSTNPEDVW 245

RESULT 6

G72300

conserved hypothetical protein - *Thermotoga maritima* (strain M3B8)

C:Species: *Thermotoga maritima*

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000

C:Accession: G72300

R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Haft, D.H.; Hickey

Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;

C.M.

Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq

A:Reference number: A72200; MUID:99287316; PMID:10360571

A:Accession: G72300

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-421 <ARN>

A:Cross-references: GB:AE001765; GB:AE000512; NID:g4981586; PIDN:AAD36122.1; PID:g498158

A:Experimental source: strain MSB

C:Genetics:

A:Gene: TM1045

C:Superfamily: arsenical pump membrane protein

Query Match

Best Local Similarity

15.3%; Score 61; DB 2; Length 421;

Matches 21; Conservative 8; Mismatches 26; Indels 18; Gaps 3;

Qy 8 FSMFMEXAGKCKMPKLIIDTPF-----STVAPALTAVLSCQLRCSLWV-----53

Db 311 FNRPLFVATVTWIS-----ILSTAFSLGAVPATLIAPTILKLVSGFPASLWVYAVGANL 366

Qy 54 GARLXPCGKVEGM 66

Db 367 GTNLTPLGAVQNI 379

A;Molecule type: DNA

R; Tamura, R.N.; Rozzo, C.; Start, L.; Chambers, J.; Reichardt, L.F.; Cooper, H.M.; Quara
J. Cell Biol. 111, 1533-1604, 1990

A; Title: Epithelial integrin alpha-6-beta-4: complete primary structure of alpha-6 and v
A; Reference number: A36429; MUID:91009492; PMID:1976638

A; Accession: A36429

A; Molecule type: mRNA

A; Residues: 1-1519, 1573-1831, 'L', 1833-1875 <TAM>

A; Cross-references: GB:X53587; NID:933950; PIDN:CAA37656.1; PID:933951

A; Experimental source: pancreatic carcinoma cell line PG

A; Note: only residues 1-96, 701-737, 1301-1519, 1573-1629, and 1833-1875 are shown

R; Kajiji, S.; Tamura, R.N.; Quaranta, V.

EMBO J. 8, 673-680, 1989

A; Title: A novel integrin (alpha(E)beta(4)) from human epithelial cells suggests a four
A; Reference number: S06962; MUID:89251596; PMID:2542022

A; Accession: S06962

A; Molecule type: protein

A; Residues: 28-42, 'Y', 44-45, 'P' <KAJ>

R; Hemler, M.E.; Crouse, C.; Sonnenberg, A.

J. Biol. Chem. 264, 6529-6535, 1989

A; Title: Association of the VLA alpha-(6) subunit with a novel protein. A possible alter
A; Reference number: A33534; MUID:89197963; PMID:2649503

A; Accession: B33534

A; Molecule type: protein

A; Residues: 28, 'XX', 31-35, 'XTXX', 40, 'X' <HEM>

C; Comment: This protein forms heterodimers with integrin alpha-6.

C; Genes: GDB:ITGB4

A; Cross-references: GDB:128028; OMIM:147557

A; Map position: 17q11-17qter

C; Superfamily: integrin beta-4 chain; EGF homology; fibronectin type III repeat homology

C; Keywords: alternative splicing; cell adhesion; chondroitin sulfate proteoglycan; cyto

F; 1-27/Domain: signal sequence #status predicted <SIG>

F; 28-1875/Product: integrin beta-4 chain, long form #status predicted <VAT>

F; 29-740/Domain: extracellular #status predicted <EXT>

F; 543-573/Domain: EGF homology <EGF>

F; 741-763/Domain: transmembrane #status predicted <TM>

F; 513/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted

Query Match 14.7%; Score 58.5; DB 2; Length 1875;
Best Local Similarity 26.9%; Pred. No. 85;
Matches 18; Conservative 10; Mismatches 30; Indels 9; Gaps 3;

QY 14 KAGKCKMFKLIIDT---PFSIVAPALTAVLSQRCRLWVGAR---LXPGG-KVE 64

DB 597 EGRCHCHQOSLYTDTICINYSATHPGLCEDLRSCVQQAQWGTGKGRICECNFKVK 656

QY 65 GMDVWRR 71

DB 657 MVDELKR 663

RESULT 12

T28915

hypothetical protein C13F10.4 - Caenorhabditis elegans

C; Species: Caenorhabditis elegans

C; Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999

C; Accession: T28915

R; Lin, A.; Wohldmann, P.

submitted to the EMBL Data Library, April 1997

A; Description: The sequence of C. elegans cosmid C13F10.

A; Reference number: Z20543

A; Accession: T28915

A; Status: preliminary;

A; Molecule type: DNA

A; Residues: 1-2076 <LIN>

A; Cross-references: EMBL:U97006; PIDN:AC47965.1; GSPDB:GN00023; CESP:C13F10.4

A; Experimental source: strain Bristol N2; clone C13F10

C; Genes: CESP:C13F10.4

A; Map position: 5

A; Introns: 114/2; 264/1; 319/3; 358/2; 664/2; 756/3; 840/3; 1022/3; 1119/3; 1840/1; 2056

Query Match 14.7%; Score 58.5; DB 2; Length 2076;

Best Local Similarity 29.9%; Pred. No. 93;
Matches 20; Conservative 11; Mismatches 17; Indels 19; Gaps 4;

QY 7 MFSMFEKA-GKCKMFKLIIDTSPSIV-----APALTAVLSQRCRLWVGARLXP 59

DB 989 MFRVFVETLSSCL-----KLLISTPTFVVVDVVGISKCLTALITC-----VGPETSC 1036

QY 60 CGKVEGM 66

DB 1037 PGVIDGV 1043

RESULT 13

T10715

MADS-box protein CMB2 - clove pink

C; Species: Dianthus caryophyllus (clove pink)

C; Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000

C; Accession: T10715

R; Baudinette, S.C.; Savin, K.W.

submitted to the EMBL Data Library, March 1995

A; Description: Carnation MADS box genes.

A; Reference number: Z17094

A; Accession: T10715

A; Status: preliminary; translated from GB/EMBL/DDBJ

A; Molecule type: mRNA

A; Residues: 1-214 <BAU>

A; Cross-references: EMBL:L40405; NID:G695318; PID:G695319

A; Experimental source: cv. Scanla; petals

C; Genes: CMB2

C; Superfamily: transcription factor squa; serum response factor DNA-binding domain homol

C; Keywords: DNA binding; transcription factor; transcription regulation

F; 2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 14.8%; Score 58; DB 2; Length 214;
Best Local Similarity 24.6%; Pred. No. 13;
Matches 14; Conservative 11; Mismatches 16; Indels 16; Gaps 2;

QY 33 IVAPALTAVLSQRCRLWVGAR-----LXP-----CGKVEGMDVWRRW 73

DB 28 IMKKAQELTVLCCDAKVSLLMISSTHKLHHVLSFGVSLKKMYDEYQKIEGVDLWRKQW 84

RESULT 14

AF3631

regulatory protein nosr [imported] - Brucella melitensis (strain 16M)

C; Species: Brucella melitensis

C; Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 21-Oct-2002

C; Accession: AF3631

R; DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova,
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A; Title: The genome sequence of the facultative intracellular pathogen Brucella melitens

A; Reference number: AD3252; PMID:11756688

A; Accession: AF3631

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-741 <KUR>

A; Cross-references: GB:AE008918; PIDN:AAL54217.1; PID:gl7985187; GSPDB:GN00191

A; Experimental source: strain 16M

C; Genes: BMEI10975

A; Map position: 11

C; Superfamily: nitrous oxide reductase expression regulator NosR

Query Match 14.6%; Score 58; DB 2; Length 741;
Best Local Similarity 35.7%; Pred. No. 41;
Matches 20; Conservative 4; Mismatches 22; Indels 10; Gaps 3;

QY 24 KLIIDTFFSIVAPALTAV-----LSQRCRLWVGARLXPCKGKVEGMDVWRRW 73

DB 592 KFRDWPBFIYVITLLAVGFVERFYCYLCPL---GAALAIPIGRIN-MPEWLKRW 643

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2004, 21:57:35 ; Search time 111.51 Seconds

(without alignments)
2443.166 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 398

Sequence: 1 ARAYXXMFMEKAGKWC.....ARLXPGKVEGMDVWRRRS 74

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xih
-Q=/cgn2_1/USPTO.spool/US10057510/runat_09032004_162236_26038/app_query.fasta_1.654
-DB=Published Applications NA -CFMT=fastap -SUFFIX=std.rnpb -MINMATCH=0.1
-LOPCL=0 -LCOEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10057510@cgn_1_1_213_@runat_09032004_162236_26038
-NCFU=6 -ICFU=3 -NO_WMAP -LARGESQRY -NEG SCORE=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5
-FGAPOPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
394	99.0	486	13	US-10-057-510-1	Sequence 1, Appli																																							
165	41.5	1415	15	US-10-131-410-47	Sequence 47, Appl																																							
165	41.5	1617	15	US-10-264-237-1023	Sequence 1023, Ap																																							
164	41.2	1076	14	US-10-062-831-17	Sequence 17, Appl																																							
164	41.2	1076	14	US-10-062-831-17	Sequence 17, Appl																																							
157	33.4	1390	14	US-10-280-953-2	Sequence 2, Appli																																							
121	30.4	499	14	US-10-280-953-5	Sequence 5, Appli																																							
95	23.9	583	14	US-10-280-953-6	Sequence 6, Appli																																							
76	19.1	2049	15	US-10-369-493-33569	Sequence 33569, A																																							
76	19.1	11945	9	US-09-922-549B-67	Sequence 67, Appl																																							
76	19.1	11945	14	US-10-114-739A-67	Sequence 67, Appl																																							
76	19.1	12728	9	US-09-922-549B-65	Sequence 65, Appl																																							
76	19.1	12728	14	US-10-251-364-6	Sequence 6, Appli																																							
76	19.1	12728	14	US-10-114-739A-65	Sequence 10, Appl																																							
76	19.1	18391	14	US-10-251-364-10	Sequence 70, Appl																																							
76	19.1	18391	14	US-10-114-739A-70	Sequence 20, Appl																																							
74.5	18.7	1967	14	US-10-083-620A-20	Sequence 20, Appl																																							
74.5	18.7	1967	14	US-10-034-015-20	Sequence 706, App																																							
74.5	18.7	3386	10	US-09-960-706-706	Sequence 446, App																																							
74.5	18.7	3386	10	US-09-873-319-446	Sequence 4, Appli																																							
74.5	18.7	3627	10	US-09-927-827-4	Sequence 1, Appli																																							
74.5	18.7	7356	10	US-09-927-827-1	Sequence 3, Appli																																							
74	18.6	1137	9	US-09-828-505-3	Sequence 1, Appli																																							
72.5	18.2	1991	15	US-10-341-434-1	Sequence 146, App																																							
72.5	18.2	1991	15	US-10-341-434-146	Sequence 2, Appli																																							
72.5	18.2	1997	9	US-09-804-741A-2	Sequence 25732, A																																							
72	18.1	1245	12	US-10-282-122A-25732	Sequence 29724, A																																							
72	18.1	1254	12	US-10-282-122A-29724	Sequence 86, Appl																																							
72	18.1	158405	14	US-10-175-523-86	Sequence 3, Appli																																							
72	18.1	158405	14	US-10-280-953-3	Sequence 163057, A																																							
71	17.8	722	15	US-10-027-632-163057	Sequence 28865, A																																							
71	17.8	722	15	US-10-027-632-163057	Sequence 105591, A																																							
71	17.8	1230	12	US-10-282-122A-28865	Sequence 3892, A																																							
70.5	17.7	146	12	US-10-425-114-32892	Sequence 3785, Ap																																							
70.5	17.7	1495	12	US-09-880-107-3785	Sequence 3, Appli																																							
70.5	17.7	4541	14	US-10-123-036-3	Sequence 73, Appl																																							
70.5	17.7	4541	14	US-10-101-510-73	Sequence 41, Appl																																							
70.5	17.7	4541	14	US-10-354-358-41	Sequence 3351, Ap																																							
70.5	17.7	4541	14	US-09-764-877-3351	Sequence 3351, Ap																																							
70	17.6	13808	15	US-10-242-515-3351	Sequence 13, Appl																																							
69.5	17.5	299	14	US-10-280-953-13	Sequence 96, Appl																																							
69.5	17.5	514	15	US-10-341-961A-96	Sequence 18022, A																																							
69.5	17.5	753	15	US-10-027-632-18022	Sequence 140287, A																																							
69.5	17.5	807	15	US-10-027-632-140287																																								

ALIGNMENTS

RESULT 1

US-10-057-510-1
; Sequence 1, Application US/10057510
; Publication No. US20020098580A1
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; APPLICANT: CuraGen Corporation
; TITLE OF INVENTION: MDN INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDN US
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: USSN 09/510,252
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USSN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USSN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 486
; TYPE: DNA

```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Any N is A, C, G, or T
US-10-057-510-1

Alignment Scores:
Pred. No.: 2,78e-47 Length: 486
Score: 394.00 Matches: 74
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.99% Indels: 0
DB: 13 Gaps: 0

US-10-057-510-2 (1-74) x US-10-057-510-1 (1-486)
QY 1 AlaArgAlaTy***LysMetPheSerMetPheMetGluLysAlaGlyLysCysTrpCys 20
Db 1 GCACGAGCGGTATNAAAAATGTTTCCATGTTTATGGAAGGCTGGGAAGTCTGGTGT 60
QY 21 LysMetProLysLeuIleAspThrProPheSerIleValAlaProAlaLeuThra 40
Db 61 AAAATGCCCAAGCTCATATAGACTACTCTTCTCCATTTGGCCCTGCTTAAGTGT 120
QY 41 ValLeuSerCysGlnLeuArgCysSerLeuTrpLeuValGlyAlaArgLeu***ProCys 60
Db 121 GTTCTTCTTGCAGCTTCTGTTTCCCTCTGGCTTGTGGGGCACGCGCTGNTCCATGT 180
QY 61 GlyLysValGluGlyMetAspValTrpArgArgArgTrpSer 74
Db 181 GGCAGGTGGGAAGGCGATGACCTGTGGAGGAGCGCTGGAGC 222

RESULT 2
US-10-131-410-47
; Sequence 47, Application US/10131410
; Publication No. US20030235915A1
; GENERAL INFORMATION:
; APPLICANT: SPECHT, THOMAS
; APPLICANT: HINZMANN, BERND
; APPLICANT: SCHMITT, ARMIN
; APPLICANT: PILARSKY, CHRISTIAN
; APPLICANT: DAHL, EDGAR
; APPLICANT: ROSENTHAL, ANDRE
; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM TISSUE OF BREAST
; FILE REFERENCE: SCH-1783
; CURRENT APPLICATION NUMBER: US/10/131,410
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 09/646,673
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: PCT/DE99/00908
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 202
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 47
; LENGTH: 1415
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-131-410-47

Alignment Scores:
Pred. No.: 1.18e-13 Length: 1415
Score: 185.00 Matches: 32
Percent Similarity: 79.55% Conservative: 3
Best Local Similarity: 72.73% Mismatches: 8
Query Match: 41.46% Indels: 1
DB: 15 Gaps: 0

US-10-057-510-2 (1-74) x US-10-131-410-47 (1-1415)
QY 32 SerIleValAlaProAlaLeuThraValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 686 TCCTTCATGCTACCCACCACCTCAGTGTGCTGAGGTCAAGGCAGCTTGTTCCTCTGG 745

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Any N is A, C, G, or T
US-10-057-510-2 (1-74) x US-10-264-237-1023 (1-1617)
QY 32 SerIleValAlaProAlaLeuThraValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 259 TCCTTCATGCTACCCACCACCTCAGTGTGCTGAGGTCAAGGCAGCTTGTTCCTCTGG 318
QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 319 CTGTGGGGGCGACGGCTGTCTCCATGTGGCAAGGTGAAGGCATGACGTGTGGAGGAG 378
QY 71 gArgTrpSer 74
Db 379 GCGCTGGAGC 388

RESULT 3
US-10-264-237-1023
; Sequence 1023, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO 1023
; LENGTH: 1617
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (92)..(92)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1539)..(1539)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1563)..(1563)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-1023

Alignment Scores:
Pred. No.: 1.41e-13 Length: 1617
Score: 165.00 Matches: 32
Percent Similarity: 79.55% Conservative: 3
Best Local Similarity: 72.73% Mismatches: 8
Query Match: 41.46% Indels: 1
DB: 15 Gaps: 0

US-10-057-510-2 (1-74) x US-10-264-237-1023 (1-1617)
QY 32 SerIleValAlaProAlaLeuThraValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 259 TCCTTCATGCTACCCACCACCTCAGTGTGCTGAGGTCAAGGCAGCTTGTTCCTCTGG 318
QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 319 CTGTGGGGGCGACGGCTGTCTCCATGTGGCAAGGTGAAGGCATGACGTGTGGAGGAG 378
QY 71 gArgTrpSer 74
Db 379 GCGCTGGAGC 388

RESULT 4
US-10-062-831-17
; Sequence 17, Application US/10062831
; Publication No. US20030105297A1
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006P1
; CURRENT APPLICATION NUMBER: US/10/062,831
```

```

; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/690,454
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: PCT/US98/10868
; PRIOR FILING DATE: May 28, 1998
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1076
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (979)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1007)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1040)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1050)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-062-831-17

Alignment Scores:
Pred. No.: 1.14e-13 Length: 1076
Score: 164.00 Matches: 33
Percent Similarity: 81.82% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 7
Query Match: 41.21% Indels: 1
DB: 14 Gaps: 0

US-10-057-510-2 (1-74) x US-10-062-831-17 (1-1076)
Qy 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 259 TCCTTCATGCTACCCACCACCTCAGTGTGAGGTCAAGGCAGCTTCGTTGTTCCCTCTGG 318
Qy 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 319 CTTGTGGGGGACCGCTGTSTYTCATGTGGCAAGGTGGAAGGCATGACGTGTGGAGGAG 378
Qy 71 gArgTrpSer 74
Db 379 GCGCTGGAGC 388

RESULT 5
US-10-062-599-17
; Sequence 17, Application US/10062599
; Publication No. US20030195346A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: PZ006P1
; CURRENT APPLICATION NUMBER: US/10/062,599
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/690,454
; PRIOR FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1076
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (979)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1007)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1040)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1050)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-062-599-17

Alignment Scores:
Pred. No.: 1.14e-13 Length: 1076
Score: 164.00 Matches: 33
Percent Similarity: 81.82% Conservative: 3
Best Local Similarity: 75.00% Mismatches: 7
Query Match: 41.21% Indels: 1
DB: 14 Gaps: 0

US-10-057-510-2 (1-74) x US-10-062-599-17 (1-1076)
Qy 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 259 TCCTTCATGCTACCCACCACCTCAGTGTGAGGTCAAGGCAGCTTCGTTGTTCCCTCTGG 318
Qy 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 319 CTTGTGGGGGACCGCTGTSTYTCATGTGGCAAGGTGGAAGGCATGACGTGTGGAGGAG 378
Qy 71 gArgTrpSer 74
Db 379 GCGCTGGAGC 388

RESULT 6
US-10-280-953-2

```

```
; Sequence 2, Application US/10280953
; Publication No. US20030113317A1
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Arvizu, Chandra S.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: MOLECULE ASSOCIATED WITH APOPTOSIS
; FILE REFERENCE: PC-0018-1 CIP
; CURRENT APPLICATION NUMBER: US/10/280,953
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 09/602,565
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/106,920
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030113317A1 2840978CBL
US-10-280-953-2

Alignment Scores:
Pred. No.: 1,64e-12 Length: 1390
Score: 157.00 Matches: 31
Percent Similarity: 77.27% Conservative: 3
Best Local Similarity: 70.45% Mismatches: 9
Query Match: 39.45% Indels: 1
DB: 14 Gaps: 0

US-10-057-510-2 (1-74) x US-10-280-953-2 (1-1390)
Qy 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 655 TCCTTCATGCTACCCACCCAGCTGCTGAGTCAAGGCGCTGTTGTTCCCTCTGG 714
Qy 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 715 CTTGTGGGGCGCGGCTGTGCTCCATGTGGCAAGGTGGAGGCGATGGACGTGTGGAGGAG 774
Qy 71 gArgTrpSer 74
Db 775 GCGCTGGAGC 784

RESULT 7
US-10-280-953-5
; Sequence 5, Application US/10280953
; Publication No. US20030113317A1
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Arvizu, Chandra S.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: MOLECULE ASSOCIATED WITH APOPTOSIS
; FILE REFERENCE: PC-0018-1 CIP
; CURRENT APPLICATION NUMBER: US/10/280,953
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 09/602,565
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/106,920
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PERL Program
; SEQ ID NO 5
; LENGTH: 499
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030113317A11441568FI

US-10-057-510-2 (1-74) x US-10-280-953-5 (1-499)
Qy 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 354 TCCTTCATGCTACCCACCCAGCTGCTGAGTCAAGGCGCTGTTGTTCCCTCTGG 413
Qy 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 414 CTTGTGGGGCGCGGCTGTGCTCCATGTGGCAAGGTGGAGGCGAT-GACGTGTGGANGANG 472
Qy 72 ArgTrp 73
Db 473 GCGTGG 478

RESULT 8
US-10-280-953-6
; Sequence 6, Application US/10280953
; Publication No. US20030113317A1
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Arvizu, Chandra S.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: MOLECULE ASSOCIATED WITH APOPTOSIS
; FILE REFERENCE: PC-0018-1 CIP
; CURRENT APPLICATION NUMBER: US/10/280,953
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: 09/602,565
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/106,920
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 583
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030113317A11441568FI
; NAME/KEY: unsure
; LOCATION: 224, 351, 379, 397, 428, 451, 502, 513, 523, 527, 529, 535, 545,
; LOCATION: 572
; OTHER INFORMATION: a, t, c, g, or other
US-10-280-953-6

Alignment Scores:
Pred. No.: 0.000446 Length: 583
Score: 95.00 Matches: 26
Percent Similarity: 75.00% Conservative: 4
Best Local Similarity: 65.00% Mismatches: 10
Query Match: 23.87% Indels: 2
DB: 14 Gaps: 0

US-10-057-510-2 (1-74) x US-10-280-953-6 (1-583)
Qy 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
```



```
Db 273 TCCTCATGCTACCCACCCTCAGTCTCAGTCTCAGGCTTGTGTTCCCTCTGG 332
Qy 52 LeuValGlyAlaArgLeu***ProCysGlyValGlyValGlyMetAspValTrpArgArg 71
Db 333 CTTGTGGGGGC-AGGCTGTTCATGTGCAAGTG-CAAGGCATGACNTGTGGAAGAGG 390

RESULT 9
US-10-369-493-33569
; Sequence 33569, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 33569
; LENGTH: 2049
; TYPE: DNA
; ORGANISM: Desulfotobacterium hafnienae
US-10-369-493-33569

Alignment Scores:
Pred. No.: 1.3 Length: 2049
Score: 76.00 Matches: 20
Percent Similarity: 48.33% Conservative: 9
Best Local Similarity: 33.33% Mismatches: 19
Query Match: 19.10% Indels: 12
DB: 15 Gaps: 3

US-10-057-510-2 (1-74) x US-10-369-493-33569 (1-2049)
Qy 9 SerMetPheMetGluLysAlaGlyLysCysTrp-----CysLysMetPro 23
Db 1226 TCATTGTAAAGAGACACAGCAAGAAAGATGGCCGAGCATGGGGCGTCCGGCTGATC 1285
Qy 24 LysLeuLeu-----IleAspThrProPheSerIleValAlaProAlaLeuThrAla 40
Db 1286 GCATTAGTCTCAAGCCGCTCCATACCATGGAGTTATTAAAGCCCTGGCGGAGCGCC 1345
Qy 41 ValLeuSerCysGlnLeuArgCysSerLeuTrpLeuValGlyAlaArgLeu***ProCys 60
Db 1346 GGGTTAAGTGCTCTATGTATGTGCACTA-----ATCCGGGCTCACTCCCTGTC 1393

RESULT 10
US-09-922-549B-67/c
; Sequence 67, Application US/09922549B
; Publication No. US20020199214A1
; GENERAL INFORMATION:
; APPLICANT: Avigenics
; TITLE OF INVENTION: Chicken Lysozyme Promoter
; FILE REFERENCE: A181 8060
; CURRENT APPLICATION NUMBER: US/09/922,549B
; CURRENT FILING DATE: 2002-04-29
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 67
; LENGTH: 11945
; TYPE: DNA
; ORGANISM: Gallus gallus
; NAME/KEY: misc_feature
; LOCATION: (1)..(237)
; OTHER INFORMATION: Sprime matrix attachment region (MAR)

US-10-057-510-2 (1-74) x US-09-922-549B-67 (1-11945)
Qy 11 PheMetGluLysAlaGlyLysCysTrpCysLysMetProLysLeuIleLeuAspThrPro 30
Db 10785 TTCACAGAGAGGGTGGTGGAGTGTCTGGCAGAGCTGCCAGAGAGGTTGTGGATCTCCA 10726
Qy 31 ----PheSerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSer 49
Db 10725 TCCTGTGAGGTGTTCATGGCCAGGTTCGATGAGCCCTGGCTGGCTGTCTAGTACCAG 10666
Qy 50 LeuTrpLeuValGlyAlaArgLeu***ProCysGlyLysValGlyMetAspValTrp 69
Db 10665 ATCTGG-----AGGTTGGTGGCCCTGC-----CTGTG 10639
Qy 69 pArgArgArgTrp 73
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (261)..(1564)
; OTHER INFORMATION: Sprime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1565)..(1912)
; OTHER INFORMATION: Sprime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1930)..(2012)
; OTHER INFORMATION: Sprime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2013)..(2671)
; OTHER INFORMATION: Intrinsically Curved DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5848)..(5934)
; OTHER INFORMATION: Transcription Enhancer
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9160)..(9325)
; OTHER INFORMATION: Transcription Enhancer
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9326)..(9626)
; OTHER INFORMATION: Negative Regulatory Element
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9621)..(9660)
; OTHER INFORMATION: Hormone Response Element
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9680)..(10060)
; OTHER INFORMATION: Hormone Response Element
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10576)..(10821)
; OTHER INFORMATION: Chicken CR1 Repeat
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10976)..(11193)
; OTHER INFORMATION: Chicken CR1 Repeat
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11424)..(11938)
; OTHER INFORMATION: Proximal promoter and lysozyme signal peptide
US-09-922-549B-67

Alignment Scores:
Pred. No.: 13.6 Length: 11945
Score: 76.00 Matches: 18
Percent Similarity: 44.62% Conservative: 11
Best Local Similarity: 27.69% Mismatches: 23
Query Match: 19.10% Indels: 13
DB: 9 Gaps: 3
```

```
Db 10638 GCGGGGGTGTGG 10626
|||||
RESULT 11
US-10-251-364-7/c
; Sequence 7, Application US/10251364
; Publication NO. US20030126629A1
; GENERAL INFORMATION:
; APPLICANT: Rapp, Jeffrey C.
; TITLE OF INVENTION: Production of a Transgenic Avian by
; TITLE OF INVENTION: Cytoplasmic Injection
; FILE REFERENCE: 11106-006-999
; CURRENT APPLICATION NUMBER: US/10/251,364
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/351,550
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/322,969
; PRIOR FILING DATE: 2001-09-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 11945
; TYPE: DNA
; ORGANISM: Gallus gallus
; NAME/KEY: misc_feature
; LOCATION: (1)..(237)
; OTHER INFORMATION: 5prime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (261)..(1564)
; OTHER INFORMATION: 5prime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1565)..(1912)
; OTHER INFORMATION: 5prime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1930)..(2012)
; OTHER INFORMATION: 5prime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2013)..(2671)
; OTHER INFORMATION: Intrinsically Curved DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5848)..(5934)
; OTHER INFORMATION: Transcription Enhancer
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9160)..(9325)
; OTHER INFORMATION: Transcription Enhancer
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9326)..(9626)
; OTHER INFORMATION: Negative Regulatory Element
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9621)..(9660)
; OTHER INFORMATION: Hormone Response Element
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9680)..(10060)
; OTHER INFORMATION: Hormone Response Element
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10576)..(10821)
; OTHER INFORMATION: Chicken CR1 Repeat
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10926)..(11193)
; OTHER INFORMATION: Chicken CR1 Repeat

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11424)..(11938)
; OTHER INFORMATION: Proximal promoter and lysozyme signal peptide
US-10-251-364-7
Alignment Scores:
Pred. No.: 13.6 Length: 11945
Score: 76.00 Matches: 18
Percent Similarity: 44.62% Conservative: 11
Best Local Similarity: 27.69% Mismatches: 23
Query Match: 19.10% Indels: 13
DB: 14 Gaps: 3
US-10-057-510-2 (1-74) x US-10-251-364-7 (1-11945)
Qy 11 PheMetGluLysAlaGlyLysCysTTPCysAlaMetProLysLeuIleLeaSpThrPro 30
Db 10785 TTCACAGAGAGGGTGTGGTCTGGCAGAGCTGCCAGAGGTTGGATGCTCCA 10726
Qy 31 ---PheSerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSer 49
Db 10725 TCCTGGAGGTGTTCATGGCCAGGTGGATGGAGCCCTGGCTGCTAGTACCAG 10666
Qy 50 LeuTrpLeuValGlyAlaArgLeu***-ProCysGlyLysValGluGlyMetAspValTr 69
Db 10665 ATCTGG-----AGGTGGTGGCCCTGC-----CTGTG 10639
Qy 69 pArgA-gArgTTP 73
Db 10638 GCGGGGGTGTGG 10626
RESULT 12
US-10-114-739A-67/c
; Sequence 67, Application US/10114739A
; Publication NO. US20030140363A1
; GENERAL INFORMATION:
; APPLICANT: Repp, Jeffrey C.
; TITLE OF INVENTION: Avian Lysozyme Promoter
; FILE REFERENCE: 11106-020-999 (AVI 013 CIP)
; CURRENT APPLICATION NUMBER: US/10/114,739A
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: 09/922,549
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: 60/280,004
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/351,550
; PRIOR FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 67
; LENGTH: 11945
; TYPE: DNA
; ORGANISM: Gallus gallus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(237)
; OTHER INFORMATION: 5prime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (261)..(1564)
; OTHER INFORMATION: 5prime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1565)..(1912)
; OTHER INFORMATION: 5prime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1930)..(2012)
; OTHER INFORMATION: 5prime matrix attachment region (MAR)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (2013)..(2671)
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: March 14, 2004, 21:56:30 ; Search time 25.0069 Seconds
(without alignments)
1642.202 Million cell updates/sec

Title: US-10-057-510-2
Perfect score: 398
Sequence: 1 ARAYXKMFMEKAGKWC.....ARLXPCGKVGMDVWRRWS 74

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/uspto spool/US10057510/runat 09032004 162235 26011/app_query.fasta_1.654
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=std.rn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10057510 -CGN 1 1 63 @runat 09032004 162235 26011 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:
5: /cgn2_6/ptodata/2/ina/PCFUS.COMB.seq:
6: /cgn2_6/ptodata/2/ina/backfiles.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	394	99.0	486	4	US-09-510-252-1
2	164	41.2	1076	4	US-09-690-454-17
3	157	39.4	1390	4	US-09-602-565-2
4	121	30.4	499	4	US-09-602-565-5
5	95	23.9	583	4	US-09-602-565-6
6	84	21.1	1326	4	US-09-489-039A-6964
7	84	21.1	2679	4	US-09-489-039A-6948
8	77	19.3	531	4	US-09-252-991A-8719
9	77	19.3	636	4	US-09-252-991A-8938
10	74.5	18.7	1967	4	US-09-256-000-20
11	74.5	18.7	3359	4	US-09-023-655-1229
12	73	18.3	2871	4	US-09-252-991A-553

C 13	72.5	18.2	1380	4	US-09-252-991A-12432	Sequence 12432, A
C 14	72.5	18.2	1620	4	US-09-252-991A-13046	Sequence 13046, A
C 15	72.5	18.2	1997	2	US-08-987-466-2	Sequence 2, Appli
C 16	72.5	18.2	1997	3	US-09-240-359-2	Sequence 2, Appli
C 17	72	18.1	771	4	US-09-252-991A-597	Sequence 597, App
C 18	72	18.1	975	4	US-09-489-039A-3630	Sequence 3630, Ap
C 19	71	17.8	260	4	US-09-602-565-3	Sequence 3, Appli
C 20	69.5	17.5	292	4	US-09-602-565-19	Sequence 19, Appl
C 21	69.5	17.5	299	4	US-09-602-565-18	Sequence 18, Appl
C 22	69.5	17.5	548	4	US-09-602-565-10	Sequence 10, Appl
C 23	69.5	17.5	577	4	US-09-602-565-11	Sequence 11, Appl
C 24	69.5	17.5	6503	4	US-09-404-650-12	Sequence 12, Appl
C 25	69.5	17.5	6503	4	US-09-935-541-12	Sequence 12, Appl
C 26	69	17.3	9713	4	US-08-961-527-43	Sequence 43, Appl
C 27	68.5	17.2	6816	4	US-09-404-650-1	Sequence 1, Appli
C 28	68.5	17.2	6816	4	US-09-935-541-1	Sequence 1, Appli
C 29	68.5	17.2	6855	4	US-09-404-650-3	Sequence 3, Appli
C 30	68.5	17.2	6855	4	US-09-935-541-3	Sequence 3, Appli
C 31	68.5	17.2	8078	4	US-09-702-251-3	Sequence 3, Appli
C 32	67	16.8	6359	4	US-09-475-252-1	Sequence 1, Appli
C 33	66.5	16.7	9837	1	US-08-832-883-68	Sequence 68, Appl
C 34	66.5	16.7	9837	2	US-08-832-877-68	Sequence 1479, Ap
C 35	66	16.6	1221	4	US-09-489-039A-1479	Sequence 3, Appli
C 36	65.5	16.5	174493	4	US-09-804-471A-3	Sequence 3, Appli
C 37	65.5	16.5	174493	4	US-10-238-709-3	Sequence 3097, Ap
C 38	64.5	16.2	2079	4	US-09-252-991A-3097	Sequence 2324, Ap
C 39	64	16.1	1275	4	US-09-252-991A-2324	Sequence 2252, Ap
C 40	64	16.1	1527	4	US-09-252-991A-2252	Sequence 516, App
C 41	64	16.1	3369	4	US-09-221-017B-516	Sequence 516, App
C 42	64	16.1	6605	1	US-08-769-309A-4	Sequence 4, Appli
C 43	64	16.1	6605	3	US-08-994-570-4	Sequence 4, Appli
C 44	64	16.1	6608	4	US-09-220-132-58	Sequence 58, Appl
C 45	63.5	16.0	1130	4	US-09-387-373-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-510-252-1
; Sequence 1, Application US/09510252
; Patent No. 6372490
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDM INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT APPLICATION NUMBER: US/09/510,252
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USSN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USSN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 486
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Any N is A, C, G, or T
US-09-510-252-1

Alignment Scores: 1.09e-43 Length: 486
Pred. No.: 394.00 Matches: 74
Score: 394.00
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.99% Indels: 0
DB: 4 Gaps: 0

US-10-057-510-2 (1-74) x US-09-510-252-1 (1-486)

QY 1 AlaArgAlaTyr***LysMetPheSerMetPheMetGluValAlaGlyLysCysTrpCys 20
DB 1 GCACGAGCGATNAAAAATGTTTCATGTTATGGAAGAGCTGGAGTGGTGGT 60
QY 21 LysMetProLysLeuIleleAspThrProPheSerIleValAlaProAlaLeuThrAla 40
DB 61 AAAATGCCCAAGCTCATAATAGTACTCTCTTCTCCATTGTTGCCCTCTCTAACTGCT 120
QY 41 ValLeuSerCysGlnLeuArgCysSerLeuTrpLeuValGlyAlaArgLeu***ProCys 60
DB 121 GTTCTTCTTTCGACGCTTCTGTTCCCTCTGCTGTTGGGCGACGGCTGTTCCATGT 180
QY 61 GlyLysValGluGlyMetAspValTrpArgArgTrpSer 74
DB 181 GCACAGGTGGAAGCATGACGTTGTGAGGAGCGCTGGAGC 222

RESULT 2

US-09-690-454-17
; Sequence 17, Application US/09690454
; Patent No. 6531447
; GENERAL INFORMATION:
; APPLICANT: Steven M. Ruben, et al.
; TITLE OF INVENTION: 32 Human Secreted Proteins
; FILE REFERENCE: P2006P1
; CURRENT APPLICATION NUMBER: US/09/690,454
; CURRENT FILING DATE: 2000-10-18
; PRIOR APPLICATION NUMBER: 09/189,144
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/044,039
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,093
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,190
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/050,935
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,101
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/048,356
; PRIOR FILING DATE: May 30, 1997
; PRIOR APPLICATION NUMBER: 60/056,250
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,296
; PRIOR FILING DATE: August 29, 1997
; PRIOR APPLICATION NUMBER: 60/056,293
; PRIOR FILING DATE: August 29, 1997
; NUMBER OF SEQ ID NOS: 229
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 1076
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (979)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1007)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1040)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1050)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-690-454-17

Alignment Scores:
Pred. No.: 1-52e-12 Length: 1076
Score: 164.00 Matches: 33
Percent Similarity: 81.82% Conservatives: 3
Best Local Similarity: 75.00% Mismatches: 7

Query Match: 41.21% Indels: 1
DB: 4 Gaps: 0
US-10-057-510-2 (1-74) x US-09-690-454-17 (1-1076)
QY 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
DB 259 TCCTTCATGCTACCCACCACTCAGTGTGAGGTCAAGGAGCTTCTGTTCTCCTCTGG 318
QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArg 71
DB 319 CTTGTGGGGGACGGCTGTCTTCATGTGGCAAGGTGAAGGCATGACGTGTGGAGGAG 378
QY 71 GArgTrpSer 74
DB 379 GCGCTGGAGC 388

RESULT 3

US-09-602-565-2
; Sequence 2, Application US/09602565
; Patent No. 6500642
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Patterson, Chandra
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: MOLECULE ASSOCIATED WITH APOPTOSIS
; FILE REFERENCE: PC-0018 US
; CURRENT APPLICATION NUMBER: US/09/602,565
; CURRENT FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/106,120
; PRIOR FILING DATE: 1998-06-29
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PERL Program
; SEQ ID NO 2
; LENGTH: 1390
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. 6500642 2840978CBI
US-09-602-565-2

Alignment Scores:

Pred. No.: 1-86e-11 Length: 1390
Score: 157.00 Matches: 31
Percent Similarity: 77.27% Conservatives: 3
Best Local Similarity: 70.45% Mismatches: 9
Query Match: 39.45% Indels: 1
DB: 4 Gaps: 0

US-10-057-510-2 (1-74) x US-09-602-565-2 (1-1390)

QY 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
DB 655 TCCTTCATGCTACCCACCACTCAGTGTGAGGTCAAGGAGCTTCTGTTCTCCTCTGG 714
QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArg 71
DB 715 CTTGTGGGGGACGGCTGTCTTCATGTGGCAAGGTGAAGGCATGACGTGTGGAGGAG 774
QY 71 GArgTrpSer 74
DB 775 GCGCTGGAGC 784

RESULT 4

US-09-602-565-5
; Sequence 5, Application US/09602565
; Patent No. 6500642
; GENERAL INFORMATION:
; APPLICANT: Yue, Henry
; APPLICANT: Patterson, Chandra
; APPLICANT: Corley, Neil C.

; APPLICANT: Guegler, Karl J.
 ; TITLE OF INVENTION: MOLECULE ASSOCIATED WITH APOPTOSIS
 ; FILE REFERENCE: PC-0018 US
 ; CURRENT APPLICATION NUMBER: US/09/602,565
 ; CURRENT FILING DATE: 2000-06-22
 ; PRIOR APPLICATION NUMBER: 09/106,120
 ; PRIOR FILING DATE: 1998-06-29
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 5
 ; LENGTH: 499
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. 6500642 2099593R6
 ; NAME/KEY: unsure
 ; LOCATION: 243, 350, 409, 449, 468, 471
 ; OTHER INFORMATION: a, t, c, g, or other
 US-09-602-565-5

Alignment Scores:

Pred. No.: 2,886-07 Length: 499
 Score: 121.00 Matches: 26
 Percent Similarity: 69.05% Conservative: 3
 Best Local Similarity: 61.90% Mismatches: 13
 Query Match: 30.40% Indels: 1
 Gaps: 4
 DB:

US-10-057-510-2 (1-74) x US-09-602-565-5 (1-499)

QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrrp 51
 DB 354 TCCTTCATGCTACCCACCACCTCAGTCTGAGGTCAGGCTGCTGTTCCCTGG 413
 QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrrpArgArg 71
 DB 414 CTTGTGGGGGACGGCTGTGTCCTGTTGGGAGGTTGAGGCAT-GACGTGTGGANGANG 472
 QY 72 ArgTrrp 73
 DB 473 GCGTGG 478

RESULT 5

US-09-602-565-6

; Sequence 6, Application US/09602565
 ; Patent No. 6500642
 ; GENERAL INFORMATION:
 ; APPLICANT: Yue, Henry
 ; APPLICANT: Patterson, Chandra
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Guegler, Karl J.
 ; TITLE OF INVENTION: MOLECULE ASSOCIATED WITH APOPTOSIS
 ; FILE REFERENCE: PC-0018 US
 ; CURRENT APPLICATION NUMBER: US/09/602,565
 ; CURRENT FILING DATE: 2000-06-22
 ; PRIOR APPLICATION NUMBER: 09/106,120
 ; PRIOR FILING DATE: 1998-06-29
 ; NUMBER OF SEQ ID NOS: 34
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 6
 ; LENGTH: 583
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No. 6500642 1441569F1
 ; NAME/KEY: unsure
 ; LOCATION: 224, 351, 379, 397, 428, 451, 502, 513, 523, 527, 529, 535, 539, 545, 572
 ; OTHER INFORMATION: a, t, c, g, or other
 US-09-602-565-6

Alignment Scores:

Pred. No.: 0.00105 Length: 583
 Score: 95.00 Matches: 26
 Percent Similarity: 75.00% Conservative: 4
 Best Local Similarity: 65.00% Mismatches: 10
 Query Match: 23.87% Indels: 2
 Gaps: 0
 DB:

US-10-057-510-2 (1-74) x US-09-602-565-6 (1-583)

QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrrp 51
 DB 273 TCCTTCATGCTACCCACCACCTCAGTCTGAGGTCAGGCTGTTGTTCCCTGG 332
 QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrrpArgArg 71
 DB 333 CTTGTGGGGGACGGCTGTGTCCTGTTGGCAAGTG-GAAGGATGACACNTGTGGAAGRG 390

RESULT 6

US-09-489-039A-6964

; Sequence 6964, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 6964
 ; LENGTH: 1326
 ; TYPE: DNA
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-6964

Alignment Scores:

Pred. No.: 0.0947 Length: 1326
 Score: 84.00 Matches: 23
 Percent Similarity: 42.42% Conservative: 5
 Best Local Similarity: 34.85% Mismatches: 24
 Query Match: 21.11% Indels: 14
 Gaps: 3
 DB:

US-10-057-510-2 (1-74) x US-09-489-039A-6964 (1-1326)

QY 8 PheSerMetPheMetGluLysAlaGlyLysCysTrpCysLysMetProLysLeuIlelle 27
 DB 410 TTCAGTATGAAGGGCGCCAGGGGCGGACGCTGCAGCTGTTCCGTACCG----- 457
 QY 28 AspThrProPheSerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArg 47
 DB 458 -----CCAAACAGTATGAATAATGCCACCGGACATCG-----CGG 493
 QY 48 CysSerLeuTrrpLeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAsp 67
 DB 494 TGC-----TGGAGGACGCGCGGCTGCCGTATCAGCTGCTGGAAGCAAGCGGTTGGCGG 547
 QY 68 ValTrrpArgArgTrrp 73
 DB 548 AAGTGGAGCGCGGCTGG 565

RESULT 7

US-09-489-039A-6948/c

; Sequence 6948, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A

;;
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 6948
LENGTH: 2679
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-6948

Alignment Scores:
Pred. No.: 0.248 Length: 2679
Score: 84.00 Matches: 23
Percent Similarity: 42.42% Conservatives: 5
Best Local Similarity: 34.85% Mismatches: 24
Query Match: 21.11% Indels: 14
DB: 4 Gaps: 3

US-10-057-510-2 (1-74) x US-09-489-039A-6948 (1-2679)

Qy 8 PheSerMetPheMetGluLysAlaGlyLysCysTrpCysLysMetProLysLeuIle 27
Db 443 TTCAAGTATGAAGGCGCCAGGCGGCGGAGCTGCGTATCGTTCGCTACG 396
Qy 28 AspThrProPheSerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArg 47
Db 395 -----CCAAACAGTATGAATGCCCGCGCACATCG-----CGG 360
Qy 48 CysSerLeuTrpLeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAsp 67
Db 359 TGC-----TGGAGGACGCGGCGTCCGTATCAGCTGCTGGAAGCGGTTGGCGG 306
Qy 68 ValTrpArgArgTrp 73
Db 305 AAGTGGAGCGCGCTGG 288

RESULT 8

US-09-252-991A-8719
Sequence 8719, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 8719
LENGTH: 531
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8719

Alignment Scores:
Pred. No.: 0.231 Length: 531
Score: 77.00 Matches: 20
Percent Similarity: 38.27% Conservatives: 11
Best Local Similarity: 24.69% Mismatches: 32
Query Match: 19.35% Indels: 18
DB: 4 Gaps: 3

US-10-057-510-2 (1-74) x US-09-252-991A-8719 (1-531)

Qy 11 PheMetGluLysAlaGlyLysCysTrp-----CysLysMet 22
Db 76 TTTCGGTGTACCGCGGCGATGCTGCGCAGATCAGTCCGCGGCGCAGCGTTCGCTGA 135
Qy 23 ProLysLeuIleAspThr-----ProPheSerIle 33

Db 136 TCAAGTGCTGTGCAAGATGCCCGAGGTTCTCAAGGTGTGCTGGCAGCCTTCTCGACG 195
Qy 34 ValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrpLeuVal 53
Db 196 ATCGCGCTTCGCGGCTGCGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCTGTCAGCG 255
Qy 54 GlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgArgTrp 73
Db 256 CCGACGCGGCTTTCACCGAGCGCGCGCTGT---TCCCTTTCGATGTGGCGCGCAGGTGG 312
Qy 74 Ser 74
Db 313 AGT 315

RESULT 9

US-09-252-991A-8938
Sequence 8938, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 8938
LENGTH: 636
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-8938

Alignment Scores:

Pred. No.: 0.296 Length: 636
Score: 77.00 Matches: 20
Percent Similarity: 38.27% Conservatives: 11
Best Local Similarity: 24.69% Mismatches: 32
Query Match: 19.35% Indels: 18
DB: 4 Gaps: 3

US-10-057-510-2 (1-74) x US-09-252-991A-8938 (1-636)

Qy 11 PheMetGluLysAlaGlyLysCysTrp-----CysLysMet 22
Db 155 TTTCGGTGTACCGCGGCGATGCTGCGCAGATCAGTCCGCGGCGCAGCGTTCGCTGA 214
Qy 23 ProLysLeuIleAspThr-----ProPheSerIle 33
Db 215 TCAAGTGCTGTGCAAGATGCCCGAGGTTCTCAAGGTGTGCTGGCAGCCTTCTCGACG 274
Qy 34 ValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrpLeuVal 53
Db 275 ATCGCGCTTCGCGGCTGCGCGTGTGCGGCGGCGGCGGCGGCGGCGGCTGTCAGCG 334
Qy 54 GlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgArgTrp 73
Db 335 CCGACGCGGCTTTCACCGAGCGCGCGCTGT---TCCCTTTCGATGTGGCGCGCAGGTGG 391
Qy 74 Ser 74
Db 392 AGT 394

RESULT 10

US-09-256-000-20/c
Sequence 20, Application US/09256000A
Patent No. 6350603
GENERAL INFORMATION:
APPLICANT: Loughney, Kate

```
; TITLE OF INVENTION: Phosphodiesterase 10
; FILE REFERENCE: 27866/35308
; CURRENT APPLICATION NUMBER: US/09/256,000A
; CURRENT FILING DATE: 1999-02-23
; EARLIER APPLICATION NUMBER: 60/075,508
; EARLIER FILING DATE: 1998-02-23
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 20
; LENGTH: 1967
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1741)
US-09-256-000-20

Alignment Scores:
Pred. No.: 3.01 Length: 1967
Score: 74.50 Matches: 22
Percent Similarity: 42.86% Conservative: 5
Best Local Similarity: 34.92% Mismatches: 28
Query Match: 18.72% Indels: 8
DB: 4 Gaps: 1

US-10-057-510-2 (1-74) x US-09-256-000-20 (1-1967)
QY 20 CysLysMetProLys-----LeuIlelleAspThrProPheSer 32
Db 309 TGCTCCAGGCCACACACCGCTGCTCCTCAGTGTCTCTCAGACAGAGTGCACGCGCTT 250
QY 33 lleValAlaProAlaLeuThraValLeuSerCysGlnLeuArgCysSerLeuTriLeu 52
Db 249 GTGGTCTCTCTGCTCCTCAGCATCAGCGAGAGTGTCTTGTGGCCACAGGCTCACATTG 190
QY 53 ValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspVal-TTPArgArgAR 72
Db 189 TACGAGGTGCGTCTTGAATTCCGGCGCATGGTGGGTGCGATGGAGACCATGGCGTGGCG 130
QY 72 GTTPSer 74
Db 129 GTGGTCA 123

RESULT 11
US-09-023-655-1229
; Sequence 1229, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
```

```
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1229:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3359 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g258761
US-09-023-655-1229

Alignment Scores:
Pred. No.: 6.26 Length: 3359
Score: 74.50 Matches: 19
Percent Similarity: 45.61% Conservative: 7
Best Local Similarity: 33.33% Mismatches: 26
Query Match: 18.72% Indels: 5
DB: 4 Gaps: 3

US-10-057-510-2 (1-74) x US-09-023-655-1229 (1-3359)
QY 17 LysCysThrCysLysMetProLysLeuIlelleAspThrProPheSerIleValAlaPro 36
Db 3031 AGGTGCTGGACCATCTCGCCAACTCACTGCCAGACGCTTCTCCGCTGCTGCCCA 3090
QY 37 AlaLeuThraValLeuSerCysGlnLeuArgCysSerLeuTriLeuValGlyAlaArg 56
Db 3091 GCC---CACGTGCTCCTTGGGTGTGCTCAGCTTCTGCTGCTCTGTTGG-----GAGAGG 3138
QY 57 Leu***ProCysGlyLysValGluGlyMetAspVal-TTPArgArgArgTTP 73
Db 3139 CAGCTGCTTGT---GTTATGTCGTGGGCCACAGTGTCTGCTGCAAGTCTGG 3186

RESULT 12
US-09-252-991A-553
; Sequence 553, Application US/09252991A
; Patent No. 8551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 553
; LENGTH: 2871
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-553

Alignment Scores:
Pred. No.: 8.01 Length: 2871
Score: 73.00 Matches: 27
Percent Similarity: 37.35% Conservative: 4
Best Local Similarity: 32.53% Mismatches: 22
Query Match: 18.34% Indels: 30
DB: 4 Gaps: 5

US-10-057-510-2 (1-74) x US-09-252-991A-553 (1-2871)
```

Qy 9 SerMetPheMetGluLysAlaGlyLysCysTrpCysLysMetProLysLeuIleLeuAsp 28
Db 2501 TCGATACGATGCTGTGCGG-----TGAATTGCGCTGGCTCATCTGGCAT 2548
Qy 29 ThrProPheSerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCys 48
Db 2549 ACCCTTTTACGCTGGCGCGCGGACTGTGCGG----- 2584
Qy 49 SerLeuTrp-----LeuValGlyAlaArgLeu***Pro----- 59
Db 2585 ACCTTTGGGGCAGCATATCGTGGGGCCGCGAGCCCGCCGAAACGACGAGG 2644
Qy 60 ---CysGlyLysValGluGlyMetAsp-----ValTrpArg 70
Db 2645 TCGTGGCGAAGCGTCAAGCGTCGACCGCAACGACGCGGGGCGCAATGTGCGGATGGAAC 2704
Qy 71 ArgArgTrp 73
Db 2705 AGGCGTGG 2713

RESULT 13

US-09-252-991A-12432/c
; Sequence 12432, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 12432
; LENGTH: 1380
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-12432

Alignment Scores:

Pred. No.: 3.42 Length: 1380
Score: 72.50 Matches: 23
Percent Similarity: 41.79% Conservative: 5
Best Local Similarity: 34.33% Mismatches: 30
Query Match: 18.22% Indels: 9
DB: 4 Gaps: 4

US-10-057-510-2 (1-74) x US-09-252-991A-12432 (1-1380)

Qy 13 GluLysAlaGlyLysCysTrpCysLysMetProLysLeuIleLeuAspThrProPheSer 32
Db 806 GAATCAGCGCGGCTGTGTCGACCGCGCGGCACTCTCAATC-----CCATGGGCT 753
Qy 33 IleValAla-----ProAlaLeuThrAlaValLeuSerCys-----GlnLeu 46
Db 752 ATTCGCGGACTGCGCCACCGCGTGTACAGCGGGCGGCGCACTGTACAGCGCTCGC 693
Qy 47 ArgCysSerLeuTrpLeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMet 66
Db 692 CGGTGCTCGCCCTGGAGCGCAGCGCGCTGTGGCGCGTGTGGCGAGTACGCGGGGGCGG 633
Qy 67 AspValTrpArgArgTrp 73
Db 632 ---TACTGGCGGAACAGGTGG 615

RESULT 14

US-09-252-991A-13046
; Sequence 13046, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 13046
; LENGTH: 1620
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-13046

Alignment Scores:

Pred. No.: 4.26 Length: 1620
Score: 72.50 Matches: 23
Percent Similarity: 41.79% Conservative: 5
Best Local Similarity: 34.33% Mismatches: 30
Query Match: 18.22% Indels: 9
DB: 4 Gaps: 4

US-10-057-510-2 (1-74) x US-09-252-991A-13046 (1-1620)

Qy 13 GluLysAlaGlyLysCysTrpCysLysMetProLysLeuIleLeuAspThrProPheSer 32
Db 839 GAATCAGCGCGGCTGTGTCGACCGCGCGGCACTCTCAATC-----CCATGGGCT 892
Qy 33 IleValAla-----ProAlaLeuThrAlaValLeuSerCys-----GlnLeu 46
Db 893 ATTCGCGGACTGCGCCACCGCGTGTACAGCGGGCGGCGCACTGTACAGCGCTCGC 952
Qy 47 ArgCysSerLeuTrpLeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMet 66
Db 953 CGGTGCTCGCCCTGGAGCGCAGCGCGCTGTGGCGCGTGTGGCGAGTACGCGGGGGCGG 1012
Qy 67 AspValTrpArgArgTrp 73
Db 1013 ---TACTGGCGGAACAGGTGG 1030

RESULT 15

US-08-987-466-2/c
; Sequence 2, Application US/08987466
; Patent No. 5922595
; GENERAL INFORMATION:
; APPLICANT: Fisher, Douglas A.
; APPLICANT: Gooding, Doug
; APPLICANT: Streeter, Dave
; TITLE OF INVENTION: CYCLIC-GMP PHOSPHODIESTERASE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Potter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/08/987,466
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; APPLICATION DATA:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.
 ; REGISTRATION NUMBER: 36,749
 ; REFERENCE/DOCKET NUMBER: PF-0442 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 650-855-0555
 ; TELEFAX: 650-845-4166
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1997 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: PROSNOT06
 ; CLONE: 828228
 ; US-08-987-466-2

Alignment Scores: Length: 1997
 Pred. No.: 5.67
 Score: 72.50
 Percent Similarity: 41.27%
 Best Local Similarity: 34.92%
 Query Match: 18.22%
 DB: 2
 Matches: 22
 Conservative: 4
 Mismatches: 29
 Indels: 8
 Gaps: 1

US-10-057-510-2 (1-74) x US-08-987-466-2 (1-1997)

```

Qy  20 CysLysMetProLys-----LeuIleIleAspThrProPheSer 32
Db  413 TGCTCCAGGCCCAACCCGCTGTCCTCAGTGGTCTCTCAGCAGACTGGCCACGGCTT 354
Qy  33 ILeValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrpLeu 52
Db  353 GTGGTTCTCTTGTCTCGACACACCGAGAGTTGCTTGTATGGCCACAGGTCTCACTTTG 294
Qy  53 ValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspVal-TipArgArgAr 72
Db  293 TACGAGTGGCTTCTGAATTCCGGGGCATGGTGGGGTCGATGGAGACCATGGCGTGTGTCG 234
Qy  72 gTrpSer 74
Db  233 GTGGTCA 227
  
```

Search completed: March 15, 2004, 00:21:06
 Job time : 29.0069 Secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2004, 21:53:34 ; Search time 122.483 Seconds
(without alignments)
2566.620 Million cell updates/sec

Title: US-10-057-510-2
Perfect score: 398
Sequence: 1 ARAYXKMFMEKAGKWC.....ARLXPCGKVGMDVWRWRS 74

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cpn2.1/USPTO.spool/US10057510/runat.09032004.162234.25965/app.query.fasta_1.654
-DB=N_Geneseq_25Jan04 -QFMT=fastap -SUFFIX=std.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10057510 @CGN 1.1 468 @runat.09032004.162234.25965 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:.*
1: Geneseqn1980s:.*
2: Geneseqn1990s:.*
3: Geneseqn2000s:.*
4: Geneseqn2001as:.*
5: Geneseqn2001bs:.*
6: Geneseqn2002as:.*
7: Geneseqn2003as:.*
8: Geneseqn2003bs:.*
9: Geneseqn2003cs:.*
10: Geneseqn2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	394	99.0	486	3 AAA75041	AAA75041 cDNA enco
2	165	41.5	1415	2 AAZ33655	Aaz33655 Human bre
3	165	41.5	1617	6 AB190461	Ab190461 Human pol
4	164	41.2	1076	2 AAV08817	Aav08817 Gene No.
5	162	40.7	1145	4 ABA08990	Abas08990 Human bre
6	157	39.4	1390	3 AAZ60460	Aaz60460 cDNA enco
7	157	39.4	1390	7 ABX94140	Abx94140 cDNA enco
8	157	39.4	1390	9 AAD61585	Aad61585 Human MAP

9	154	38.7	1480	4 AAK94527	Aak94527 Human ful
10	121	30.4	499	3 AAZ60479	Aaz60479 Partial h
11	121	30.4	499	7 ABX94143	Abx94143 Incyte ID
12	121	30.4	499	9 AAD61588	Aad61588 Human MAP
13	95	23.9	583	3 AAZ60480	Aaz60480 Partial h
14	95	23.9	583	7 ABX94144	Abx94144 Incyte ID
15	95	23.9	583	9 AAD61589	Aad61589 Human MAP
16	78	19.6	3809	4 ABL10069	Ab110069 Drosophil
17	78	19.6	7080	4 ABL10068	Ab110068 Drosophil
18	76.5	19.2	1401	4 ABL17985	Ab117985 Drosophil
19	76	19.1	11945	7 ACC69399	Acc69399 Chicken l
20	76	19.1	11945	7 AAD48553	Aad48553 Chicken l
21	76	19.1	11945	8 ACC69384	Acc69384 Lysozyme
22	76	19.1	12728	7 ACC69398	Acc69398 Chimeric
23	76	19.1	12728	7 AAD48557	Aad48557 Chimeric
24	76	19.1	12728	8 ACC69363	Acc69363 Lysozyme
25	76	19.1	18391	7 ACC69402	Acc69402 Chicken l
26	76	19.1	18391	8 ACC69367	Acc69367 Chicken l
27	76	19.1	21329	7 AAD48562	Aad48562 Chimeric
28	75	18.8	3459	4 ABL15746	Ab115746 Drosophil
29	75	18.8	5160	4 ABL28041	Ab128041 Drosophil
30	75	18.8	7671	4 ABL28040	Ab128040 Drosophil
31	74.5	18.7	1967	2 AAZ06763	Aaz06763 Phosphodi
32	74.5	18.7	3386	6 ABK84112	Abk84112 Human cDN
33	74.5	18.7	3386	6 ABK64551	Abk64551 Human ben
34	74.5	18.7	3627	9 ADD24876	Add24876 Xanthomon
35	74.5	18.7	7356	9 ADD24942	Add24942 Xanthomon
36	74	18.6	1137	6 AAS15594	Aas15594 Gene vacc
37	73.5	18.5	877	4 ABL14853	Ab114853 Drosophil
38	73.5	18.5	3007	4 ABL14852	Ab114852 Drosophil
39	72.5	18.2	1185	7 ABZ21533	Abz21533 Humulus l
40	72.5	18.2	1300	7 ABZ21534	Abz21534 Humulus l
41	72.5	18.2	1991	6 ABK92161	Abk92161 Prostata
42	72.5	18.2	1997	2 AAZ09086	Aaz09086 Cyclic-GM
43	72	18.1	1245	7 ACA37862	Aca37862 Prokaryot
44	72	18.1	1254	7 ACA41854	Aca41854 Prokaryot
45	72	18.1	33962	3 AAA81528	Aaa81528 N. mening

ALIGNMENTS

RESULT 1

AAA75041
ID AAA75041 standard; cDNA; 486 BP.

XX AC AAA75041;

XX DT 02-JAN-2001 (first entry)

XX DE cDNA encoding a human MDM2 interacting polypeptide (MDMIP).

XX KW Human, MDM2 interacting polypeptide; MDMIP; MDM2; cell cycle progression; cell differentiation; cancer; sarcoma; glioma; squamous cell carcinoma; breast cancer; astrocytoma; leukemia; lymphoma; tumorigenesis; gene therapy; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 1..225

XX FT /*tag= a

XX FT /product= "MDM2 interacting polypeptide (MDMIP)"

XX FT /transl_except= (pos: 13..15, aa: Xaa)

XX FT /transl_except= (pos: 172..174, aa: Xaa)

XX FT /note= "Xaa is an unknown amino acid; partial sequence"

XX XX WO2000050590-A1.

XX PD 31-AUG-2000.

XX XX 23-FEB-2000; 2000WO-US004582.

XX XX 23-FEB-1999; 99US-0121192P.

PR 03-MAR-1999; 99US-0122643P.
 PR 22-FEB-2000; 2000US-00510252.
 XX (CURA-) CURAGEN CORP.
 XX Nandabalan K, Yang M, Schulz VP;
 XX WPI; 2000-558398/51.
 DR P-PSDB; AAB08845.
 XX Novel MDM2 interacting protein useful for treating or preventing
 PT disorders involving aberrant levels of MDM2 and/or MDM-interacting
 PT proteins, comprises a specific amino acid sequence.
 XX Claim 24; Fig 1; 78pp; English.
 XX The present sequence encodes a human MDM2 interacting polypeptide
 CC (MDMIP). MDMIP was identified using a yeast two hybrid system, using a
 CC fragment of MDM2 as the bait protein. The MDMIP polypeptide is useful for
 CC detecting and removing MDM2 polypeptides in a biological sample by
 CC forming MDM2-MDMIP complexes. MDMIP and MDM2 are useful to identify
 CC compounds or other agents which modulate the activity of MDM2 and/or
 CC MDMIP-mediated processes. Agents that modulate the function of MDMIP/MDM2
 CC complexes are useful for treating and preventing a disease or disorder
 CC involving aberrant levels of MDM2 or MDMIP. MDMIP is also useful for
 CC treating diseases caused by aberrant levels of expression of MDM2 genes,
 CC such as disorders of cell cycle progression, cell differentiation, and
 CC transcriptional control, including cancers such as human sarcoma, glioma,
 CC squamous cell carcinoma, breast cancer, astrocytoma, leukemia and
 CC lymphoma, and tumorigenesis. MDMIP and MDM2 nucleic acids are useful in
 CC gene therapy
 XX SQ Sequence 486 BP; 104 A; 113 C; 145 G; 114 T; 0 U; 10 Other;

Alignment Scores:
 Pred. No.: 4 64e-40 Length: 486
 Score: 394.00 Matches: 74
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.99% Indels: 0
 DB: 3 Gaps: 0

US-10-057-510-2 (1-74) x AAA75041 (1-486)
 QY 1 AlaArgAlaTyrr***LysMetPheSerMetPheMetGluLysAlaGlyLysCysTrpCys 20
 Db 1 GCACGAGCGGTATNAAAAATGTTTCCATGTTTATGGAAGGCTGGGAAGTGGTGT 60
 QY 21 LysMetProLysLeuIleAapThrProPheSerIleValAlaProAlaLeuThra 40
 Db 61 AAAATGCCAGCTCATATAGATACTCTTCTCCATTTGTCCTGCTCTAACTGCT 120
 QY 41 ValLeuSerCysGlnLeuArgCysSerLeuValGlyAlaArgLeu***ProCys 60
 Db 121 GTTCTTCTTCGACGCTGCTTGTTCCTCTGCTGCTGGGGACGCGGTGNTCCATGT 180
 QY 61 GlyLysValGluGlyMetAspValTrpArgArgArgTrpSer 74
 Db 181 GGCAAGGTGGAGGACATGACGTGTGGAGGAGCGCTGGAGC 222

RESULT 2
 AAZ33655
 ID AAZ33655 standard; cDNA; 1415 BP.
 XX
 AC AAZ33655;
 XX
 XX 08-DEC-1999 (first entry)
 DT
 XX Human breast tumour-associated EST 45.
 DE
 XX Expressed sequence tag; EST; human; breast; cancer; gene therapy;
 KW treatment; tumour; cytostatic; medicament; ss.
 XX

OS Homo sapiens.
 XX DE19813839-A1.
 XX 23-SEP-1999.
 XX 20-MAR-1998; 98DE-01013839.
 XX 20-MAR-1998; 98DE-01013839.
 XX (META-) METAGEN GES GENOMFORSCHUNG MBH.
 XX Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E, Rosenthal A;
 XX WPI; 1999-528981/45.
 XX Human nucleic acid sequences and protein products from tumor breast
 PT tissue, useful for breast cancer therapy.
 XX Claim 1a; 124; 188pp; German.
 XX This invention describes novel human nucleic acid sequences from tumor
 CC breast tissue which have cytostatic activity. The nucleic acid sequences
 CC can be used to produce and isolate full-length gene sequences. They can
 CC be used to express proteins, which can be used as tools to find an
 CC activity against breast cancer. The sequences can be used in sense or
 CC antisense form. They are especially useful for medicaments for gene
 CC therapy to treat breast cancer. AAZ33611-248617 represents expressed
 CC sequence tags described in the method of the invention
 XX SQ Sequence 1415 BP; 246 A; 439 C; 502 G; 228 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 9 44e-11 Length: 1415
 Score: 165.00 Matches: 32
 Percent Similarity: 79.55% Conservative: 3
 Best Local Similarity: 72.73% Mismatches: 8
 Query Match: 41.46% Indels: 1
 DB: 2 Gaps: 0

US-10-057-510-2 (1-74) x AAZ33655 (1-1415)
 QY 32 SerIleValAlaProAlaLeuThraAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 Db 686 TCCTTCATGCTACCCACCACCTCAGTCTGAGGTCAAGGACGCTTCGTTGTTCCCTCTGG 745
 QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArg 71
 Db 746 CTTCTGGGGCACGGCTGTGCTCCATGTGCAAGGTGGAAGCATGCATGTGGAGGAG 805
 QY 71 GArgTrpSer 74
 Db 806 GCGCTGGAGC 815

RESULT 3
 ABL90461
 ID ABL90461 standard; cDNA; 1617 BP.
 XX
 AC ABL90461;
 XX
 DT 24-MAY-2002 (first entry)
 XX Human polynucleotide SEQ ID NO 1023.
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antinflammatory; antitumor;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein; gene; ss.
 XX
 OS Homo sapiens.
 XX WO200190304-A2.

XX 29-NOV-2001.
 PD
 XX
 PF 18-MAY-2001; 2001WO-US016450.
 XX
 PF 19-MAY-2000; 2000US-02055515P.
 PR
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 PI Birse CE, Rosen CA;
 PI
 XX
 DR WPI; 2002-122018/16.
 DR F-PSDB; ABB90052.
 XX
 PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders.
 XX
 PS Claim 4; SEQ ID NO 1023; 2081pp + Sequence Listing; English.
 CC
 CC The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1617 BP; 311 A; 493 C; 482 G; 316 T; 0 U; 15 Other;
 Alignment Scores:
 Pred. No.: 1-12e-10 Length: 1617
 Score: 165.00 Matches: 32
 Percent Similarity: 79.55% Conservative: 3
 Best Local Similarity: 72.73% Mismatches: 8
 Query Match: 41.46% Indels: 1
 DB: 6 Gaps: 0
 US-10-057-510-2 (1-74) x ABL90461 (1-1617)
 QY 32 SerileValAlaProAlaLeuThraValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 DB 259 TCCTTCATGCTACCCACCACCTCAGTGTGAGGTCAGGTCAGTCTGCTGCTGCTGCTG 318
 QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArg 71
 DB 319 CTTGTGGGGCAGCGCTGTGCTCCATGTGGCAGGTCAGGTCAGTGTGAGGTCAGGTCAG 378
 QY 71 GArgTrpSer 74
 DB 379 GCGTGGAGC 388
 RESULT 4
 AAV08817
 ID AAV08817 standard; cDNA; 1076 BP.
 XX
 AC AAV08817;
 XX
 XX 19-FEB-1999 (first entry)
 DT
 XX
 DE Gene No. 7 encoding human secreted protein.
 XX

KW Secreted protein; human; protein therapy; gene therapy; blood disorder;
 KW pathological condition; diagnosis; cancer; neurological disorder;
 KW developmental abnormality; foetal deficiency; leukaemia; hepatic disease;
 KW immune system disorder; Alzheimer's disease; cognitive disorder;
 KW schizophrenia; prostate disease; autoimmune disorder; AIDS; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO9854206-A1.
 XX
 PD 03-DEC-1998.
 XX
 PF 28-MAY-1998; 98WC-US010868.
 XX
 XX 30-MAY-1997; 97US-0044039P.
 PR 30-MAY-1997; 97US-0048093P.
 PR 30-MAY-1997; 97US-0048101P.
 PR 30-MAY-1997; 97US-0048190P.
 PR 30-MAY-1997; 97US-0048356P.
 PR 30-MAY-1997; 97US-0050935P.
 PR 29-AUG-1997; 97US-0056250P.
 PR 29-AUG-1997; 97US-0056293P.
 PR 29-AUG-1997; 97US-0056296P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ruben SM, Rosen CA, Carter KC, Dillon PJ, Endress GA, Yu G;
 PI Ni J, Peng P;
 PI WPI; 1999-070209/06.
 XX P-PSDB; AAW73403.
 DR
 XX New isolated human genes - useful for diagnosis and treatment of, e.g.
 PT cancers, neurological disorders, immune diseases, developmental disorders
 PT or blood disorders.
 XX
 PS Claim 1; Page 116-117; 188pp; English.
 XX
 CC This sequence represents a cDNA of the invention, and is designated Gene
 CC No. 7. This sequence encodes a human secreted protein, and is expressed
 CC in various human tissues including macrophages. The DNA sequences of the
 CC invention and their corresponding secreted polypeptides are useful for
 CC preventing, treating or ameliorating medical conditions, e.g. by protein
 CC or gene therapy. Also pathological conditions can be diagnosed by
 CC determining the amount of the new polypeptides in a sample or by
 CC ascertaining the presence of mutations in the DNA sequences. Specific uses
 CC are described for each of the DNA sequences and the encoded proteins,
 CC based on which tissues they are most highly expressed in, and include
 CC developing products for the diagnosis or treatment of cancer, tumours,
 CC neurological disorders, developmental abnormalities and foetal
 CC deficiencies, blood disorders, leukaemias, diseases of the immune system
 CC (including allergies or asthma), hepatic disease, Alzheimer's and
 CC cognitive disorders, schizophrenia, prostate diseases, autoimmune
 CC disorders and AIDS. The polypeptides are also useful for identifying
 CC their binding partners
 XX
 SQ Sequence 1076 BP; 214 A; 320 C; 332 G; 202 T; 0 U; 8 Other;
 Alignment Scores:
 Pred. No.: 8-83e-11 Length: 1076
 Score: 164.00 Matches: 33
 Percent Similarity: 81.82% Conservative: 3
 Best Local Similarity: 75.00% Mismatches: 7
 Query Match: 41.21% Indels: 1
 DB: 2 Gaps: 0
 US-10-057-510-2 (1-74) x AAV08817 (1-1076)

QY 32 SerileValAlaProAlaLeuThraValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 DB 259 TCCTTCATGCTACCCACCACCTCAGTGTGAGGTCAGGTCAGTCTGCTGCTGCTG 318
 QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArg 71

Db 319 CTTGTGGGACCGCTGCTGTCATGTCGACAGGTGGAGGACGTGTGGAGGAG 378
QY 71 gArgTrpSer 74
Db 379 GCGCTGGAGC 388
RESULT 5
ABAO8990
ID ABAO8990 standard; cDNA; 1145 BP.
AC ABAO8990;
XX
XX
DT 11-JAN-2002 (first entry)
DE Human breast tumour-associated protein homologue cDNA, SEQ ID NO:766.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerrary; antilucer; ss.
XX
OS Homo sapiens.
XX
XX WO2001571188-A2.
XX
PD 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US003800.
XX
XX 03-FEB-2000; 2000US-00496914.
XX
XX 27-APR-2000; 2000US-00560875.
XX
XX (HYSB-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-457746/49.
XX
XX P-PSDB; ABB11746.
XX
XX Human proteins and DNA encoding sequences useful for preventing, treating
XX or ameliorating a medical condition in a mammalian subject e.g. arthritis
XX and cancer.
XX
XX Claim 1; Page 685; 1963pp; English.
XX
XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
XX sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
XX invention also relates to vectors and recombinant host cells comprising a
XX nucleotide of the invention, methods of producing the novel polypeptides,
XX antibodies against the polypeptides, methods of detecting the nucleotides
XX or polypeptides in a sample, and methods of identifying compounds which
XX bind to polypeptides of the invention. Although novel, many of the
XX polypeptides of the invention have homology to known proteins, thereby
XX giving an insight into their probable biological activities, and hence
XX potential therapeutic applications. The polypeptides of the invention may
XX have various activities, including cytokine, cell proliferation or cell
XX differentiation activities; stem cell growth factor activity;
XX haematopoiesis regulatory activity; tissue growth activity;
XX immunomodulatory activity; activin- or inhibin-related activities;
XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
XX thrombolytic activities; receptor or ligand activities; or may be
XX involved in oncogenesis, cancer cell proliferation or metastasis.
XX Depending on their biological activities, polypeptides and nucleotides of

the invention are useful for preventing, treating or ameliorating medical
conditions, e.g., by protein or gene therapy. Such conditions include
cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
proliferative retinopathy, atherosclerosis, coronary heart disease,
arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
vascular growth. Polypeptides involved with tissue regeneration and
repair (or nucleic acids encoding them) may be used to promote wound
healing (e.g., of burns, incisions and ulcers), while those with
immunomodulatory activities may be used in the treatment of viral,
bacterial and fungal infections in addition to immune disorders.
Polypeptides with growth factor activity may be used in cell cultures to
promote cell growth. For example, such polypeptides may be used to
manipulate stem cells in culture to give rise to neuroepithelial cells
that can be used to augment or replace cells damaged by illness,
autoimmune disease or accidental damage. The polypeptides and nucleotides
may also be used in the diagnosis of the above conditions, and in drug
screening techniques. The present sequence represents a cDNA encoding a
novel human polypeptide of the invention

Sequence 1145 BP; 216 A; 355 C; 372 G; 202 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.71e-10 Length: 1145
Score: 162.00 Matches: 31
Percent Similarity: 75.55% Conservative: 4
Best Local Similarity: 70.45% Mismatches: 8
Query Match: 40.70% Indels: 1
DB: 4 Gaps: 0

US-10-057-510-2 (1-74) x ABA08990 (1-1145)

QY 32 SerIleValAlaProAlaLeuthrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 406 TCCTTCATGCTACCCACCACCTCAGTGTGAGGCAAGGAGCTTCGTGTTCCTCTGG 465
QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 466 CTTGTGGGGGACCGCTGTGCTTCATGTGTCAGAGTGGAGGATGCACGTGTGGAGGAG 525
QY 71 gArgTrpSer 74
Db 526 GCGCTGGAGC 535

RESULT 6

AAZ60460

ID AAZ60460 standard; cDNA; 1390 BP.

XX

AC AAZ60460;

XX

DT 05-MAY-2000 (first entry)

XX

DE cDNA encoding a human molecule associated with apoptosis 3 (MAPOP-3).

XX

XX Human; molecule associated with apoptosis; MAPOP; MAPOP-1; MAPOP-2;

XX MAPOP-3; cell proliferative disorder; arteriosclerosis; cirrhosis;

XX psoriasis; cancer; autoimmune disease; rheumatoid arthritis; allergy;

XX acquired immune deficiency syndrome; Crohn's disease; infection; ss.

OS Homo sapiens.

XX

Key Location/Qualifiers

CDS 198..833

/*tag= a

/product= "MAPOP-2"

XX

XX WO200000609-A2.

XX

PD 06-JAN-2000.

XX

XX 23-JUN-1999; 99WO-US014188.

XX

XX 29-JUN-1998; 98US-00106920.

PR

XX PA (INCY-) INCYTE PHARM INC.
XX PI Hillman JL, Yue H, Tang YT, Corley NC, Guegler KJ, Patterson C;
XX PR WPI; 2000-170917/15.
XX DR P-PSDB; AAY68680.
XX PA
XX PT New proteins associated with apoptosis, used for diagnosis, treatment and
XX PT prevention of cell proliferative and autoimmune disorders.
XX PS
XX PS Claim 7; Page 61; 72pp; English.
XX CC The present sequence encodes a human molecule associated with apoptosis 3
XX CC (MAPOP-3). The specification also describes MAPOP-2 and MAPOP-1. The
XX CC MAPOP-3 nucleic acids were first identified in Incyte clone 2840978 from
XX CC a dorsal root ganglion cDNA library. The MAPOP-3 cDNA sequence is a
XX CC consensus sequence derived from the overlapping and/or extended sequences
XX CC given in AA260477-82. The MAPOP polynucleotides and polypeptides are used
XX CC in human or veterinary medicine, to treat or prevent cell proliferative
XX CC disorders (e.g. arteriosclerosis, cirrhosis, psoriasis or many forms of
XX CC cancer), and for treating or preventing autoimmune diseases (e.g.
XX CC acquired immune deficiency syndrome, allergy, Crohn's disease, rheumatoid
XX CC arthritis, infection etc.). The polypeptides are also used to screen for
XX CC specific antagonists or agonists, which are potentially useful as
XX CC therapeutic agents. Antagonists are used to treat autoimmune diseases
XX SQ Sequence 1390 BP; 247 A; 437 C; 481 G; 225 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: Length: 1390
Score: 157.00 Matches: 31
Percent Similarity: 77.27% Conservative: 3
Best Local Similarity: 70.45% Mismatches: 9
Query Match: 39.45% Indels: 1
DB: Gaps: 0

US-10-057-510-2 (1-74) x AA260460 (1-1390)
QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 655 TCCTTCATGCTACCCACCCTCAGTCTGAGTCAAGGACGCTTTGTTCTCCTCTGG 714
QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 715 CTTGTGGGGGCACGGCTGTGCTCCATGTGGCAAGGTGGAAGGCATGACGTGTGGAGGAG 774
QY 71 gAgTTrpSer 74
Db 775 GCGCTGGAGC 784

RESULT 7
ABX94140
ID ABX94140 standard; cDNA; 1390 BP.
AC ABX94140;
XX
XX 06-JUN-2003 (first entry)
XX cDNA encoding human molecule associated with apoptosis, MAPOP-3.
XX Human; T-cell death associated gene; TDAG; MAPOP-3; gene expression;
KW molecule associated apoptosis; mammalian model system; breast cancer;
KW breast adenocarcinoma; cytostatic; Incyte ID No 2840978CD1; Gene; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 251..634
XX /*tag= a
XX /*product= "MAPOP-3"
XX
XX US6500642-B1.

XX PD 31-DEC-2002.
XX PF 22-JUN-2000; 2000US-00602565.
XX PR 29-JUN-1998; 98US-00106920.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Yue H, Patterson C, Corley NC, Guegler KJ;
XX PS WPI; 2003-352181/33.
XX DR P-PSDB; AB008391.
XX PT New isolated cDNA encoding a T-cell death associated polypeptide, MAPOP-
XX PT 3, useful for diagnosing and treating breast adenocarcinoma.
XX PS Claim 2; Fig 1; 34pp; English.
XX CC The present invention relates to the isolation of a novel mammalian T-
XX CC cell death associated gene (TDAG) protein or molecule associated with
XX CC apoptosis, designated MAPOP-3. Also disclosed are polynucleotide
XX CC sequences encoding it. A polynucleotide sequence encoding MAPOP-3 is
XX CC useful for producing the protein, for detecting differential gene
XX CC expression, for screening a library or several molecules or compounds to
XX CC identify at least one ligand that specifically binds to the
XX CC polynucleotide sequence, and for producing a mammalian model system. The
XX CC MAPOP-3 polypeptide and polynucleotide sequences are useful in the
XX CC diagnosis and treatment of breast adenocarcinoma. The polynucleotide
XX CC sequence encoding MAPOP-3 is useful for producing transgenic cell lines
XX CC or organisms which are model systems for human breast cancer, and upon
XX CC which the toxicity and efficacy of potential treatments may be tested.
XX CC The present sequence encoding human MAPOP-3 is isolated from an Incyte
XX CC cDNA clone (Incyte ID No:2840978CD1)
XX SQ Sequence 1390 BP; 247 A; 437 C; 481 G; 225 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: Length: 1390
Score: 157.00 Matches: 31
Percent Similarity: 77.27% Conservative: 3
Best Local Similarity: 70.45% Mismatches: 9
Query Match: 39.45% Indels: 1
DB: Gaps: 0

US-10-057-510-2 (1-74) x ABX94140 (1-1390)
QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 655 TCCTTCATGCTACCCACCCTCAGTCTGAGTCAAGGACGCTTTGTTCTCCTCTGG 714
QY 52 LeuValGlyAlaArgLeu-***ProCysGlyLysValGluGlyMetAspValTrpArgAr 71
Db 715 CTTGTGGGGGCACGGCTGTGCTCCATGTGGCAAGGTGGAAGGCATGACGTGTGGAGGAG 774
QY 71 gAgTTrpSer 74
Db 775 GCGCTGGAGC 784

RESULT 8
AAD61585
ID AAD61585 standard; cDNA; 1390 BP.
XX
XX AC AAD61585;
XX
XX 15-JAN-2004 (first entry)
XX Human MAPOP-3 cDNA.
XX
XX DE
XX Human; T cell death-associated protein; MAPOP-3; diagnosis; therapy;
KW cancer; breast adenocarcinoma; Gene; ss.
XX
XX Homo sapiens.

```
XX PH Key Location/Qualifiers
XX CDS 251..634
XX FT /*tag= a
XX TT /product= "Human MAPOP-3 protein"
XX PN US2003113317-A1.
XX PD 19-JUN-2003.
XX PF 23-OCT-2002; 2002US-00280953.
XX PR 29-JUN-1998; 98US-00106920.
XX PR 22-JUN-2000; 2000US-00602565.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Yue H, Arvizu CS, Corley NC, Guegler KJ;
XX DR WPI; 2003-810875/76.
XX DR P-PSDB; ABW00797.
XX PT Novel purified T cell death-associated protein MAPOP-3 useful for
XX PT diagnosing, monitoring or treating cancers such as breast adenocarcinoma.
XX PS Disclosure; Page 22-23; Opp; English.
XX CC The present invention relates to novel T cell death-associated proteins
XX CC (MAPOP-3) and polynucleotides encoding such proteins. Sequences of the
XX CC invention are useful for diagnosing, monitoring or treating cancers such
XX CC as breast adenocarcinoma. The present sequence is human MAPOP-3 cDNA
XX SQ Sequence 1390 BP; 247 A; 437 C; 481 G; 225 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9,288-10 Length: 1390
Score: 157.00 Matches: 31
Percent Similarity: 77.27% Conservative: 3
Best Local Similarity: 70.45% Mismatches: 9
Query Match: 39.45% Indels: 1
DB: Gaps: 0

US-10-057-510-2 (1-74) x AAD51595 (1-1390)
QY 32 SerLeValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 655 TCCTTCATGCTACCCACCACCTCAGTGTGAGTCAAGGCAGCTTTGTTGTTCCCTCTGG 714
QY 52 LeuValGlyAlaArgLeu-***ProCysGlyValGluGlyMetAspValTrpArgAr 71
Db 715 CTGTGGGGGACCGCTGTGCTCCATGTGGCAAGGTGGAGCATGACGTGTGGAGGAG 774
QY 71 GArgTrpSer 74
Db 775 GCGCTGGAGC 784

RESULT 9
AAK94527
ID AAK94527 standard; cDNA; 1480 BP.
XX AC AAK94527;
XX DT 06-NOV-2001 (first entry)
XX DE Human full-length cDNA, SEQ ID NO: 3400.
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX OS Homo sapiens.
XX PN BP130094-A2.
XX OS 05-SEP-2001.
```

```
XX 07-JUL-2000; 2000EP-00114089.
XX PF
XX PR 08-JUL-1999; 99JP-00194486.
XX PR 11-JAN-2000; 2000JP-00118774.
XX PR 02-MAY-2000; 2000JP-00183765.
XX XX
XX PA (HELI-) HELIX RES INST.
XX XX
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX DR WPI; 2001-524255/58.
XX DR P-PSDB; AAM93595.
XX PT 830 Primers useful for synthesizing full length cDNA clones and their use
XX PT in genetic manipulation.
XX XX
XX PS Claim 8; SEQ ID NO 3400; 1380pp + Sequence Listing; English.
XX CC The invention relates to primers for synthesising full length cDNA
XX CC clones. 830 cDNA molecules encoding a human protein have been isolated
XX CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
XX CC been determined. Primers for synthesising the full length cDNA are useful
XX CC for clarifying the function of the protein encoded by the cDNA. The full
XX CC length clones were obtained by construction of full length enriched cDNA
XX CC libraries that were synthesised by the oligo-capping method. The primers
XX CC enable the production of the full length cDNA easily without any special
XX CC methods. The present sequence is a full length human cDNA of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in CD-ROM format directly
XX CC from EPO
XX SQ Sequence 1480 BP; 257 A; 457 C; 534 G; 232 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,398-09 Length: 1480
Score: 154.00 Matches: 30
Percent Similarity: 77.27% Conservative: 4
Best Local Similarity: 68.18% Mismatches: 9
Query Match: 38.69% Indels: 1
DB: Gaps: 0

US-10-057-510-2 (1-74) x AAK94527 (1-1480)
QY 32 SerLeValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 753 TCCTTCATGCTACCCACCACCTCAGTGTGAGGCAAGGCAGCTTTGTTGTTCCCTCTGG 812
QY 52 LeuValGlyAlaArgLeu-***ProCysGlyValGluGlyMetAspValTrpArgAr 71
Db 813 CTGTGGGGGACCGCTGTGCTCCATGTGGCAAGGTGGAGCATGACGTGTGGAGGAG 872
QY 71 GArgTrpSer 74
Db 873 GCGCTGGAGC 882

RESULT 10
AAZ60479
ID AAZ60479 standard; cDNA; 499 BP.
XX AC AAZ60479;
XX DT 05-MAY-2000 (first entry)
XX DE Partial human molecule associated with apoptosis 3 (MAPOP-3) cDNA.
XX KW Human; molecule associated with apoptosis; MAPOP; MAPOP-1; MAPOP-2;
XX KW MAPOP-3; cell proliferative disorder; arteriosclerosis; cirrhosis;
XX KW psoriasis; cancer; autoimmune disease; rheumatoid arthritis; allergy;
XX KW acquired immune deficiency syndrome; Crohn's disease; infection; ss.
XX OS Homo sapiens.
```

XX WO200000609-A2.
XX PD 06-JAN-2000.
XX PF 23-JUN-1999; 99WO-US014188.
XX PR 29-JUN-1998; 98US-00106920.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Hillman JL, Yue H, Tang YT, Corley NC, Guegler KJ, Patterson C;
XX WPI; 2000-170917/15.
XX New proteins associated with apoptosis, used for diagnosis, treatment and
XX prevention of cell proliferative and autoimmune disorders.
XX Disclosure; Page 69; 72pp; English.
XX AA260477-82 represent overlapping and/or extended sequences that were
XX used to produce a consensus sequence (AA260460) encoding a human molecule
XX associated with apoptosis (MAPOP-3). The specification also describes
XX MAPOP-1 and MAPOP-2. The MAPOP-3 nucleic acids were first identified in
XX Incyte clone 2840978 from a dorsal root ganglion cDNA library. The MAPOP
XX polynucleotides and polypeptides are used in human or veterinary
XX medicine, to treat or prevent cell proliferative disorders (e.g.
XX arteriosclerosis, cirrhosis, psoriasis or many forms of cancer), and for
XX treating or preventing autoimmune diseases (e.g. acquired immune
XX deficiency syndrome, allergy, Crohn's disease, rheumatoid arthritis,
XX infection etc.). The polypeptides are also used to screen for specific
XX antagonists or agonists, which are potentially useful as therapeutic
XX agents. Antagonists are used to treat autoimmune diseases
XX
SQ Sequence 499 BP; 94 A; 151 C; 170 G; 78 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 8,02e-06 Length: 499
Score: 121.00 Matches: 26
Percent Similarity: 69.05% Conservative: 3
Best Local Similarity: 61.90% Mismatches: 13
Query Match: 30.40% Indels: 1
DB: 3 Gaps: 0

US-10-057-510-2 (1-74) x AA260479 (1-499)
QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 354 TCCTTCATGCTACCCACCACCTCAGTGTGAGGTCAAGGCAGCTTCGTTGCCNCTGG 413
QY 52 LeuValGlyAlaArgLeu***ProCysGlyIysValGluGlyMetAspValTrpArgArg 71
Db 414 CTTGTGGGGGACCGGCTGTGTCCATGTGGGAAGGTNGAAGGCAT-GACGTGTGGANGANG 472
QY 72 ArgTrp 73
Db 473 GCGTGG 478
RESULT 11
ABX94143
ID ABX94143 standard; cDNA; 499 BP.
XX AC ABX94143;
XX
DT 06-JUN-2003 (first entry)
XX
XX Incyte ID No 2099593R6 cDNA clone used to obtain human MAPOP-3 cDNA.
XX Human; T-cell death associated gene; TDAG; MAPOP-3; Gene expression;
XX molecule associated apoptosis; mammalian model system; breast cancer;
XX breast adenocarcinoma; cytostatic; Incyte ID No 2099593R6; ss.
XX
XX Homo sapiens.

XX US6500642-B1.
XX PD 31-DEC-2002.
XX PF 22-JUN-2000; 2000US-00602565.
XX PR 29-JUN-1998; 98US-00106920.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Yue H, Patterson C, Corley NC, Guegler KJ;
XX WPI; 2003-352181/33.
XX New isolated cDNA encoding a T-cell death associated polypeptide, MAPOP-
XX 3, useful for diagnosing and treating breast adenocarcinoma.
XX Disclosure; Col 35-36; 34pp; English.
XX The present invention relates to the isolation of a novel mammalian T-
XX cell death associated gene (TDAG) protein or molecule associated with
XX apoptosis, designated MAPOP-3. Also disclosed are polynucleotide
XX sequences encoding it. A polynucleotide sequence encoding MAPOP-3 is
XX useful for producing the protein, for detecting differential gene
XX expression, for screening a library or several molecules or compounds to
XX identify at least one ligand that specifically binds to the
XX polynucleotide sequence, and for producing a mammalian model system. The
XX MAPOP-3 polypeptide and polynucleotide sequences are useful in the
XX diagnosis and treatment of breast adenocarcinoma. The polynucleotide
XX sequence encoding MAPOP-3 is useful for producing transgenic cell lines
XX or organisms which are model systems for human breast cancer, and upon
XX which the toxicity and efficacy of potential treatments may be tested.
XX ABX94141-ABX94146 represent incyte cDNA clones used to obtain cDNA
XX encoding human MAPOP-3
XX
SQ Sequence 499 BP; 94 A; 151 C; 170 G; 78 T; 0 U; 6 Other;

Alignment Scores:
Pred. No.: 8,02e-06 Length: 499
Score: 121.00 Matches: 26
Percent Similarity: 69.05% Conservative: 3
Best Local Similarity: 61.90% Mismatches: 13
Query Match: 30.40% Indels: 1
DB: 3 Gaps: 0

US-10-057-510-2 (1-74) x ABX94143 (1-499)
QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
Db 354 TCCTTCATGCTACCCACCACCTCAGTGTGAGGTCAAGGCAGCTTCGTTGCCNCTGG 413
QY 52 LeuValGlyAlaArgLeu***ProCysGlyIysValGluGlyMetAspValTrpArgArg 71
Db 414 CTTGTGGGGGACCGGCTGTGTCCATGTGGGAAGGTNGAAGGCAT-GACGTGTGGANGANG 472
QY 72 ArgTrp 73
Db 473 GCGTGG 478
RESULT 12
AAD61588
ID AAD61588 standard; cDNA; 499 BP.
XX AC AAD61588;
XX
DT 15-JAN-2004 (first entry)
XX
XX Human MAPOP-3 cDNA fragment #3.
XX Human; T cell death-associated protein; MAPOP-3; diagnosis; therapy;
XX cancer; breast adenocarcinoma; ss.
XX

OS Homo sapiens.
 XX US2003113317-A1.
 PN
 XX
 XX 19-JUN-2003.
 XD
 XX
 XX 23-OCT-2002; 2002US-00280953.
 PP
 XX
 XX 29-JUN-1998; 98US-00106920.
 PR
 XX 22-JUN-2000; 2000US-00602565.
 PR
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Yue H, Arvizu CS, Corley NC, Guegler KJ;
 PI
 XX WPI; 2003-810875/76.
 DR
 XX
 XX Novel purified T cell death-associated protein MAPOP-3 useful for
 PT diagnosing, monitoring or treating cancers such as breast adenocarcinoma.
 PT
 XX Disclosure; Page 24; Opp; English.
 XX
 XX The present invention relates to novel T cell death-associated proteins
 CC (MAPOP-3) and polynucleotides encoding such proteins. Sequences of the
 CC invention are useful for diagnosing, monitoring or treating cancers such
 CC as breast adenocarcinoma. The present sequence is human MAPOP-3 cDNA
 CC fragment
 XX
 XX Sequence 499 BP; 94 A; 151 C; 170 G; 78 T; 0 U; 6 Other;
 SQ
 Alignment Scores:
 Pred. No.: 8,02e-06 Length: 499
 Score: 121.00 Matches: 26
 Percent Similarity: 69.05% Conservative: 3
 Best Local Similarity: 61.90% Mismatches: 13
 Query Match: 30.40% Indels: 1
 DB: 9 Gaps: 0
 US-10-057-510-2 (1-74) x AAD61588 (1-499)
 QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 Db 354 TCCTTCATGCTACCCACCTCAGTCTGAGTCAAGGCAGCTTCGTTGCCNCTGG 413
 QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgArg 71
 Db 414 CTGTGGGGGCAGGGCTGTGTCCATGTGGGAAGGTGAAGGCAT-CAGCTGTGGANGANG 472
 QY 72 ArgTrp 73
 Db 473 GCGTGG 478
 RESULT 13
 AAZ60480
 ID AAZ60480 standard; cDNA; 583 BP.
 XX
 XX AAZ60480;
 AC
 XX
 XX 05-MAY-2000 (first entry)
 DT
 XX
 XX Partial human molecule associated with apoptosis 3 (MAPOP-3) cDNA.
 DE
 XX Human; molecule associated with apoptosis; MAPOP-1; MAPOP-2;
 KW MAPOP-3; cell proliferative disorder; arteriosclerosis; cirrhosis;
 KW psoriasis; cancer; autoimmune disease; rheumatoid arthritis; allergy;
 KW acquired immune deficiency syndrome; Crohn's disease; infection; ss.
 XX
 XX Homo sapiens.
 OS
 XX WC200000609-A2.
 PN
 XX 06-JAN-2000.
 PD
 XX

PF 23-JUN-1999; 99WO-US014188.
 XX
 XX 29-JUN-1998; 98US-00106920.
 PR
 XX (INCY-) INCYTE PHARM INC.
 PA
 XX Hillman JL, Yue H, Tang YT, Corley NC, Guegler KJ, Patterson C;
 PI
 XX WPI; 2000-170917/15.
 DR
 XX
 XX New proteins associated with apoptosis, used for diagnosis, treatment and
 PT prevention of cell proliferative and autoimmune disorders.
 PT
 XX Disclosure; Page 69-70; 72pp; English.
 XX
 XX AAZ60477-82 represent overlapping and/or extended sequences that were
 CC used to produce a consensus sequence (AAZ60460) encoding a human molecule
 CC associated with apoptosis 3 (MAPOP-3). The specification also describes
 CC MAPOP-1 and MAPOP-2. The MAPOP-3 nucleic acids were first identified in
 CC Incyte clone 2840978 from a dorsal root ganglion cDNA library. The MAPOP
 CC polynucleotides and polypeptides are used in human or veterinary
 CC medicine, to treat or prevent cell proliferative disorders (e.g.
 CC arteriosclerosis, cirrhosis, psoriasis or many forms of cancer), and for
 CC treating or preventing autoimmune diseases (e.g. acquired immune
 CC deficiency syndrome, allergy, Crohn's disease, rheumatoid arthritis,
 CC infection etc.). The polypeptides are also used to screen for specific
 CC antagonists or agonists, which are potentially useful as therapeutic
 CC agents. Antagonists are used to treat autoimmune diseases
 XX
 XX Sequence 583 BP; 125 A; 163 C; 185 G; 95 T; 0 U; 15 Other;
 SQ
 Alignment Scores:
 Pred. No.: 0.0178 Length: 583
 Score: 95.00 Matches: 26
 Percent Similarity: 75.00% Conservative: 4
 Best Local Similarity: 65.00% Mismatches: 10
 Query Match: 23.87% Indels: 2
 DB: 3 Gaps: 0
 US-10-057-510-2 (1-74) x AAZ60480 (1-583)
 QY 32 SerileValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTrp 51
 Db 273 TCCTTCATGCTACCCACCTCAGTCTGAGTCAAGGCAGCTTCGTTGCCNCTGG 332
 QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTrpArgArg 71
 Db 333 CTGTGGGGGC-AGGCTGTNTCCATGTGGCAGTGTG-CAAGGCATGCACNTGTGGAGAGG 390
 RESULT 14
 ABX94144
 ID ABX94144 standard; cDNA; 583 BP.
 XX
 XX ABX94144;
 AC
 XX
 XX 06-JUN-2003 (first entry)
 DT
 XX
 XX Incyte ID No 1441568F1 cDNA clone used to obtain human MAPOP-3 cDNA.
 DE
 XX Human; T-cell death associated gene; TDAG; MAPOP-3; gene expression;
 KW molecule associated apoptosis; mammalian model system; breast cancer;
 KW breast adenocarcinoma; cytostatic; Incyte ID No 1441568F1; ss.
 XX
 XX Homo sapiens.
 OS
 XX US6500642-B1.
 PN
 XX 31-DEC-2002.
 PD
 XX
 XX 22-JUN-2000; 2000US-00602565.
 PP
 XX 29-JUN-1998; 98US-00106920.
 PR
 XX

PA (INCY-) INCYTE GENOMICS INC.
XX
PI Yue H, Patterson C, Corley NC, Guegler KJ;
XX
DR WPI; 2003-352181/33.
XX
XX New isolated cDNA encoding a T-cell death associated polypeptide, MAPOP-
PT 3, useful for diagnosing and treating breast adenocarcinoma.
PT
XX
PS Disclosure; Col 35-36; 34pp; English.
XX
XX The present invention relates to the isolation of a novel mammalian T-
CC cell death associated gene (TPAG) protein or molecule associated with
CC apoptosis, designated MAPOP-3. Also disclosed are polynucleotide
CC sequences encoding it. A polynucleotide sequence encoding MAPOP-3 is
CC useful for producing the protein, for detecting differential gene
CC expression, for screening a library or several molecules or compounds to
CC identify at least one ligand that specifically binds to the
CC polynucleotide sequence, and for producing a mammalian model system. The
CC MAPOP-3 polypeptide and polynucleotide sequences are useful in the
CC diagnosis and treatment of breast adenocarcinoma. The polynucleotide
CC sequence encoding MAPOP-3 is useful for producing transgenic cell lines
CC or organisms which are model systems for human breast cancer, and upon
CC which the toxicity and efficacy of potential treatments may be tested.
CC ABX94141-ABX94146 represent incyte cDNA clones used to obtain cDNA
CC encoding human MAPOP-3
XX
SQ Sequence 583 BP; 125 A; 163 C; 185 G; 95 T; 0 U; 15 Other;
XX
Alignment Scores:
Pred. No.: 0.0178 Length: 583
Score: 95.00 Matches: 26
Percent Similarity: 75.00% Conservative: 4
Best Local Similarity: 65.00% Mismatches: 10
Query Match: 23.87% Indels: 2
DB: 7 Gaps: 0
US-10-057-510-2 (1-74) x ABX94144 (1-583)
QY 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTTP 51
Db 273 TCCTTCATGCTACCCACACCTCAGTGTGAGGTCAGGAGCTTCGTTGCTCCTCTGG 332
QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTTPArgArg 71
Db 333 CTTGTGGGGGC-AGGCTGTNTCCATGTGGCAAGTG-GAAGGCATGACNTGTGGAGAGG 390
RESULT 15
AAD61589
ID AAD61589 standard; cDNA; 583 BP.
XX
AC AAD61589;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human MAPOP-3 cDNA fragment #4.
XX
KW Human; T cell death-associated protein; MAPOP-3; diagnosis; therapy;
KW cancer; breast adenocarcinoma; ss.
XX
OS Homo sapiens.
XX
PN US2003113317-A1.
XX
PD 19-JUN-2003.
XX
PF 23-OCT-2002; 2002US-00280953.
XX
PR 29-JUN-1998; 98US-00106920.
PR 22-JUN-2000; 2000US-00602565.
XX
PA (INCY-) INCYTE GENOMICS INC.

PI Yue H, Arvizu CS, Corley NC, Guegler KJ;
XX
DR WPI; 2003-810875/76.
XX
PT Novel purified T cell death-associated protein MAPOP-3 useful for
PT diagnosing, monitoring or treating cancers such as breast adenocarcinoma.
PT
XX
PS Disclosure; Page 24; Opp; English.
XX
XX The present invention relates to novel T cell death-associated proteins
CC (MAPOP-3) and polynucleotides encoding such proteins. Sequences of the
CC invention are useful for diagnosing, monitoring or treating cancers such
CC as breast adenocarcinoma. The present sequence is human MAPOP-3 cDNA
CC fragment
XX
SQ Sequence 583 BP; 125 A; 163 C; 185 G; 95 T; 0 U; 15 Other;
XX
Alignment Scores:
Pred. No.: 0.0178 Length: 583
Score: 95.00 Matches: 26
Percent Similarity: 75.00% Conservative: 4
Best Local Similarity: 65.00% Mismatches: 10
Query Match: 23.87% Indels: 2
DB: 9 Gaps: 0
US-10-057-510-2 (1-74) x AAD61589 (1-583)
QY 32 SerIleValAlaProAlaLeuThrAlaValLeuSerCysGlnLeuArgCysSerLeuTTP 51
Db 273 TCCTTCATGCTACCCACACCTCAGTGTGAGGTCAGGAGCTTCGTTGCTCCTCTGG 332
QY 52 LeuValGlyAlaArgLeu***ProCysGlyLysValGluGlyMetAspValTTPArgArg 71
Db 333 CTTGTGGGGGC-AGGCTGTNTCCATGTGGCAAGTG-GAAGGCATGACNTGTGGAGAGG 390

Search completed: March 14, 2004, 22:05:55
Job time : 126.483 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:45:03 ; Search time 19.6483 Seconds
(without alignments)
795.253 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 398

Sequence: 1 ARAYXKMFMEKAGKWC.....ARLXPGKVGMDVWRRWS 74

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	394	99.0	74	US-10-057-510-2	Sequence 2, Appl
2	69	17.3	1032	US-09-733-643-16	Sequence 16, Appl
3	69	17.3	1032	US-10-120-801-64	Sequence 1673, Ap
4	65.5	16.5	142	US-10-094-749-1973	Sequence 1973, Ap
5	65.5	16.5	383	US-09-759-130B-410	Sequence 410, Appl
6	65.5	16.5	383	US-10-042-431-40	Sequence 40, Appl
7	61.5	15.5	330	US-10-295-027-1366	Sequence 1366, Ap
8	61.5	15.5	504	US-10-104-047-3872	Sequence 3872, Ap
9	61.5	15.5	703	US-10-108-260A-4545	Sequence 4545, Ap
10	61.5	15.5	760	US-10-295-027-1367	Sequence 1367, Ap
11	60.5	15.2	439	US-09-922-501-12	Sequence 12, Appl
12	60.5	15.2	439	US-09-557-796-14	Sequence 14, Appl
13	60	15.1	514	US-10-369-493-8445	Sequence 8445, Ap
14	58.5	14.7	224	US-09-989-920-274	Sequence 274, Appl
15	58.5	14.7	224	US-10-465-572-4	Sequence 4, Appl

16	58.5	14.7	1752	15	US-10-116-275-205	Sequence 205, App
17	58.5	14.7	1752	15	US-10-295-027-360	Sequence 360, App
18	58.5	14.7	1822	15	US-10-295-027-1254	Sequence 1254, Ap
19	58	14.6	130	9	US-09-908-805B-29	Sequence 29, Appl
20	57.5	14.4	338	13	US-10-042-417-12	Sequence 12, Appl
21	57.5	14.4	509	14	US-10-156-761-7900	Sequence 7900, Ap
22	56	14.1	194	14	US-10-017-161-262	Sequence 262, App
23	56	14.1	194	15	US-10-292-798-232	Sequence 232, App
24	56	14.1	553	14	US-10-225-567A-520	Sequence 520, App
25	56	14.1	654	15	US-10-369-493-5059	Sequence 5059, Ap
26	56	14.1	772	9	US-09-909-320-339	Sequence 339, App
27	56	14.1	772	9	US-09-909-088B-339	Sequence 339, App
28	56	14.1	772	9	US-09-905-291A-339	Sequence 339, App
29	56	14.1	772	9	US-09-902-853-339	Sequence 339, App
30	56	14.1	772	9	US-09-907-824-339	Sequence 339, App
31	56	14.1	772	9	US-09-907-841-339	Sequence 339, App
32	56	14.1	772	10	US-09-904-011-339	Sequence 339, App
33	56	14.1	772	10	US-09-906-742-339	Sequence 339, App
34	56	14.1	772	10	US-09-906-838-339	Sequence 339, App
35	56	14.1	772	10	US-09-907-613-339	Sequence 339, App
36	56	14.1	772	10	US-09-907-942-339	Sequence 339, App
37	56	14.1	772	10	US-09-904-859-339	Sequence 339, App
38	56	14.1	772	10	US-09-909-204-339	Sequence 339, App
39	56	14.1	772	10	US-09-904-820-339	Sequence 339, App
40	56	14.1	772	10	US-09-904-786-339	Sequence 339, App
41	56	14.1	772	10	US-09-906-646-339	Sequence 339, App
42	56	14.1	772	10	US-09-906-700-339	Sequence 339, App
43	56	14.1	772	10	US-09-903-786-339	Sequence 339, App
44	56	14.1	772	10	US-09-902-903-339	Sequence 339, App
45	56	14.1	772	10	US-09-903-749A-339	Sequence 339, App

ALIGNMENTS

RESULT 1

US-10-057-510-2
; Sequence 2, Application US/10057510
; Publication No. US20020098580A1
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDM INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT APPLICATION NUMBER: US/10/057,510
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: USSN 09/510,252
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USSN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USSN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Any X can be any amino acid.
US-10-057-510-2

Query Match 99.0%; Score 394; DB 13; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.7e-41;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAYXKMFMEKAGKWCMPKLIIDTPFSIVAPALTAVLSQIRCSLWLVGARLXPC 60
DB 1 ARAYXKMFMEKAGKWCMPKLIIDTPFSIVAPALTAVLSQIRCSLWLVGARLXPC 60
QY 61 GKVEGDVWRRWS 74


```
RESULT 5
US-09-759-130B-410
; Sequence 410, Application US/09759130B
; Publication No. US2003022279A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: McCarthy, Sean A
; APPLICANT: Fraser, Christopher C
; APPLICANT: Sharp, John D
; APPLICANT: Barnes, Thomas S
; APPLICANT: Kirst, Susan J
; APPLICANT: Mackay, Charles R
; APPLICANT: Myers, Paul S
; APPLICANT: Leiby, Kevin R
; APPLICANT: Wrighton, Nicolas
; APPLICANT: Goodearl, Andrew
; APPLICANT: Holtzman, Douglas A
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC, AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: MPI00-5350NMIM
; CURRENT APPLICATION NUMBER: US/09/759,130B
; PRIOR APPLICATION NUMBER: US 09/479,249
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: US 09/559,497
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/596,194
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/342,364
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: US 09/608,452
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/393,996
; PRIOR FILING DATE: 1999-09-10
; PRIOR APPLICATION NUMBER: US 09/602,871
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 09/420,707
; PRIOR FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 460
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 410
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-759-130B-410

Query Match 16.5%; Score 65.5; DB 10; Length 383;
Best Local Similarity 30.2%; Pred. No. 6.8;
Matches 13; Conservative 8; Mismatches 19; Indels 3; Gaps 1;

Qy 18 CWCMPKLIIDTPFSIVAPALTAVLSQCLRLVGLVGLXPC 60
Db 185 CWCLSVLMFLPHNAYKSIATGISCILACLVILL---LSPC 224

RESULT 6
US-10-042-431-40
; Sequence 40, Application US/10042431
; Publication No. US20020182675A1
; GENERAL INFORMATION:
; APPLICANT: MCCARTHY, Sean A
; APPLICANT: BARNES, Thomas M
; APPLICANT: FRASER, Christopher C
; APPLICANT: SHARP, John D
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING DIAGNOSTIC,
; TITLE OF INVENTION: PREVENTIVE, THERAPEUTIC, AND OTHER USES
; FILE REFERENCE: 10147-602
```

```
; CURRENT APPLICATION NUMBER: US/10/042,431
; CURRENT FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 09/333,159
; PRIOR FILING DATE: 1999-06-14
; PRIOR APPLICATION NUMBER: US 09/578,063
; PRIOR FILING DATE: 2000-05-24
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 40
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-042-431-40

Query Match 16.5%; Score 65.5; DB 13; Length 383;
Best Local Similarity 30.2%; Pred. No. 6.8;
Matches 13; Conservative 8; Mismatches 19; Indels 3; Gaps 1;

Qy 18 CWCMPKLIIDTPFSIVAPALTAVLSQCLRLVGLVGLXPC 60
Db 185 CWCLSVLMFLPHNAYKSIATGISCILACLVILL---LSPC 224

RESULT 7
US-10-295-027-1366
; Sequence 1366, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1366
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1366

Query Match 15.5%; Score 61.5; DB 15; Length 330;
```


Best Local Similarity 28.2%; Pred. No. 18;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;

QY 4 YXKMFSPM--EKAGKWCCKMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXP-- 59
Db 134 YSKNISLMNMFQPPSKAW-RASQMTFFIFLLFFPSFTGVL-CTLAITW----RLKPSA 187
QY 60 -CGKVEGMDVW 69
Db 188 DCGPFRGLPLF 198

RESULT 8

US-10-104-047-3872
; Sequence 3872, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20030236392A1e1 full length cDNA
; FILE REFERENCE: H1-A0105
; CURRENT APPLICATION NUMBER: US/10/104,047
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3872
; LENGTH: 504
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-104-047-3872

Query Match 15.5%; Score 61.5; DB 15; Length 504;
Best Local Similarity 28.2%; Pred. No. 28;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;

QY 4 YXKMFSPM--EKAGKWCCKMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXP-- 59
Db 308 YSKNISLMNMFQPPSKAW-RASQMTFFIFLLFFPSFTGVL-CTLAITW----RLKPSA 361
QY 60 -CGKVEGMDVW 69
Db 362 DCGPFRGLPLF 372

RESULT 9

US-10-108-260A-4545
; Sequence 4545, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4545
; LENGTH: 703
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-4545

Query Match 15.5%; Score 61.5; DB 15; Length 703;
Best Local Similarity 28.2%; Pred. No. 40;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;

QY 4 YXKMFSPM--EKAGKWCCKMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXP-- 59
Db 564 YSKNISLMNMFQPPSKAW-RASQMTFFIFLLFFPSFTGVL-CTLAITW----RLKPSA 617
QY 60 -CGKVEGMDVW 69
||| : : :

Db 618 DCGPFRGLPLF 628

RESULT 10

US-10-295-027-1367
; Sequence 1367, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1367
; LENGTH: 760
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1367

Query Match 15.5%; Score 61.5; DB 15; Length 760;
Best Local Similarity 28.2%; Pred. No. 44;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;

QY 4 YXKMFSPM--EKAGKWCCKMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXP-- 59
Db 564 YSKNISLMNMFQPPSKAW-RASQMTFFIFLLFFPSFTGVL-CTLAITW----RLKPSA 617
QY 60 -CGKVEGMDVW 69
Db 618 DCGPFRGLPLF 628

RESULT 11

US-09-922-501-12
; Sequence 12, Application US/09922501
; Patent No. US20020120119A1
; GENERAL INFORMATION:
; APPLICANT: Dartois, Veronique A.
; APPLICANT: Hoch, James A.
; APPLICANT: Valle, Fernando

```

; APPLICANT: Kumar, Manoj
; TITLE OF INVENTION: 2, 5-DKG PERMEASES
; FILE REFERENCE: P-SR 4877
; CURRENT APPLICATION NUMBER: US/09/922,501
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/633,294
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 09/677,032
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Klebsiella oxytoca
US-09-922-501-12

Query Match
Best Local Similarity 15.2%; Score 60.5; DB 9; Length 439;
Matches 18; Conservative 8; Mismatches 22; Indels 7; Gaps 2;

QY 7 MFSMFMEKAGKWCWKMPKLIIDTPFSIVAPALTAVLSQRLCSLWLVGARLXPCG 61
Db 296 LFSSLSDRTKG-----RKLFCVCLP--LIGFALCMFLSVALKNQIWLVSYAALVCGC 343

RESULT 12
US-09-557-796-14
; Sequence 14, Application US/09557796
; Publication No. US20030073140A1
; GENERAL INFORMATION:
; APPLICANT: Hoch, James
; APPLICANT: Dattois, Veronique
; TITLE OF INVENTION: METABOLIC SELECTION METHODS
; FILE REFERENCE: 234/191
; CURRENT APPLICATION NUMBER: US/09/557,796
; CURRENT FILING DATE: 2000-04-25
; PRIOR FILING DATE: 1998-10-14
; PRIOR APPLICATION NUMBER: 09/172,952
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Yiax2
US-09-557-796-14

Query Match
Best Local Similarity 15.2%; Score 60.5; DB 10; Length 439;
Matches 18; Conservative 8; Mismatches 22; Indels 7; Gaps 2;

QY 7 MFSMFMEKAGKWCWKMPKLIIDTPFSIVAPALTAVLSQRLCSLWLVGARLXPCG 61
Db 296 LFSSLSDRTKG-----RKLFCVCLP--LIGFALCMFLSVALKNQIWLVSYAALVCGC 343

RESULT 13
US-10-369-493-8445
; Sequence 8445, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
```

```

; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 8445
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Ralstonia metallidurans
US-10-369-493-8445
```

```

Query Match
Best Local Similarity 15.1%; Score 60; DB 15; Length 514;
Matches 19; Conservative 7; Mismatches 17; Indels 12; Gaps 4;

QY 27 IDTPFSI-VAPALT--AVLSQRL-----CSLWLVGARLXPCGKVEGMDVWR 70
Db 3 VDTPLPIGTAPALPVHASNMCVRAPEQLPQCRCWPNANVL-AGVIDGRNIWR 56
```

```

RESULT 14
US-09-989-920-274
; Sequence 274, Application US/09989920
; Patent No. US20020172957A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 60/252,500
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 274
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-274
```

```

Query Match
Best Local Similarity 14.7%; Score 58.5; DB 9; Length 224;
Matches 15; Conservative 14; Mismatches 14; Indels 9; Gaps 2;

QY 21 KMPKLIIDTPFSIVAPALTAVLSQRLCSLWLVGARLXPCGKVEGMDVWR 72
Db 8 ELFENILLELFTHV-PAROLLNCLVCSLWR-----DLIDLVTWKRK 50
```

```

RESULT 15
US-10-465-572-4
; Sequence 4, Application US/10465572
; Publication No. US20030207840A1
; GENERAL INFORMATION:
; APPLICANT: Riggins, Gregory
; APPLICANT: Lai, Anita
; TITLE OF INVENTION: GENES INDUCED BY HYPOXIA
; FILE REFERENCE: 000250.00012
; CURRENT APPLICATION NUMBER: US/10/465,572
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US/10/201,642
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/307,600
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 224
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-465-572-4
```

```

Query Match
14.7%; Score 58.5; DB 15; Length 224;
```

Best Local Similarity 28.8%; Pred. No. 29;
Matches 15; Conservative 14; Mismatches 14; Indels 9; Gaps 2;

Qy 21 KMPKLIIDTPFSIVAPALTAVLSCQLRCSLWLVGARLXPCGKVEGMDVWRR 72

Db 8 ELPENILLEFTHV-PARQLLNCLVCSLWR-----DLIDLVTLWKRK 50

Search completed: March 14, 2004, 21:51:19
Job time : 20.6483 secs

GenCore version 5.1.6
Copyright (c) 1995 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:42:43 ; Search time 9.69655 Seconds
(without alignments)
393.988 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 398

Sequence: 1 ARAYXXMFMEKAGKCMC.....ARLXPGKVGMDVRRWS 74

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
 - 2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
 - 3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
 - 4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
 - 5: /cgn2_6/ptodata/2/iaa/PTCUS.COMB.pep.*
 - 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	394	99.0	74	4	US-09-510-252-2
2	77	19.3	176	4	US-09-252-991A-25290
3	62.5	15.7	331	4	US-09-489-039A-12478
4	61	15.3	240	4	US-09-489-039A-12478
5	60.5	15.2	439	4	US-09-172-952-14
6	60	15.1	480	4	US-09-252-991A-25416
7	58	14.6	130	4	US-09-302-769-29
8	57.5	14.4	341	4	US-09-252-991A-29845
9	57	14.3	371	4	US-09-252-991A-21750
10	56.5	14.2	720	4	US-09-252-991A-23212
11	56	14.1	553	3	US-08-845-546-12
12	56	14.1	772	4	US-09-907-794A-339
13	56	14.1	772	4	US-09-905-125A-339
14	56	14.1	772	4	US-09-902-775A-339
15	55	13.8	141	4	US-09-252-991A-29876
16	54.5	13.7	203	4	US-09-252-991A-32870
17	54	13.6	133	4	US-09-252-991A-28869
18	54	13.6	179	2	US-08-845-539-6
19	54	13.6	179	4	US-09-362-642-6
20	54	13.6	251	4	US-09-134-000C-4324
21	53.5	13.4	168	4	US-09-252-991A-32502
22	53.5	13.4	204	4	US-09-252-991A-17837
23	53	13.3	69	4	US-09-345-236B-7
24	53	13.3	300	4	US-09-252-991A-27295
25	53	13.3	346	1	US-08-119-773-5
26	53	13.3	357	1	US-08-119-773-2
27	53	13.3	357	1	US-08-119-773-4

28	53	13.3	357	1	US-08-119-773-6	Sequence 6, Appli
29	53	13.3	357	4	US-09-874-132-27	Sequence 27, Appli
30	53	13.3	414	4	US-09-252-991A-31151	Sequence 31151, A
31	53	13.3	429	4	US-09-489-039A-13801	Sequence 13801, A
32	53	13.3	553	4	US-09-252-991A-17089	Sequence 17089, A
33	52.5	13.2	95	4	US-09-543-681A-5593	Sequence 5593, Ap
34	52.5	13.2	148	4	US-09-252-991A-19469	Sequence 19469, A
35	52.5	13.2	360	4	US-09-252-991A-22238	Sequence 22238, A
36	52.5	13.2	403	4	US-09-543-681A-7076	Sequence 7076, Ap
37	52.5	13.2	421	4	US-09-543-681A-5230	Sequence 5230, Ap
38	52.5	13.2	456	4	US-09-252-991A-28459	Sequence 28459, A
39	52.5	13.2	728	4	US-08-219-237B-7	Sequence 7, Appli
40	52	13.1	162	2	US-08-477-347-16	Sequence 16, Appli
41	52	13.1	162	3	US-08-476-862-7	Sequence 7, Appli
42	52	13.1	162	3	US-08-468-560C-7	Sequence 7, Appli
43	52	13.1	162	3	US-09-800-909-7	Sequence 7, Appli
44	52	13.1	162	4	US-09-800-908-16	Sequence 16, Appli
45	52	13.1	162	4	US-09-800-908-16	Sequence 16, Appli

ALIGNMENTS

RESULT 1

US-09-510-252-2
; Sequence 2, Application US/09510252
; Patent No. 6372490
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDM INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT APPLICATION NUMBER: US/09/510,252
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USSN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USSN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Any X can be any amino acid.
US-09-510-252-2

Query Match 99.0%; Score 394; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 9.1e-45;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ARAYXXMFMEKAGKCMKPKLIIDTPFSIVAPALTAVLSCQLRCSLWLGARLXPC 60
Db 1 ARAYXXMFMEKAGKCMKPKLIIDTPFSIVAPALTAVLSCQLRCSLWLGARLXPC 60
Qy 61 GKVEGMDVRRWS 74
Db 61 GKVEGMDVRRWS 74

RESULT 2

US-09-252-991A-25290
; Sequence 25290, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25290
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25290

Query Match
Best Local Similarity 19.3%; Score 77; DB 4; Length 176;
Matches 20; Conservative 11; Mismatches 32; Indels 18; Gaps 3;

QY 11 FMEKAGKWC-----CKMPKLIIDT-----PFSIVAPALTAVLSQCLRCSLWLV 53
DB 26 FVSGRCWRSSPGACRRSRCARSPVSRRCRWQPSSTIAPSPVSCRRNRASAWSA 85
QY 54 GARLXPCGKVEGMDVWRWRS 74
DB 86 PTALSPAARC-SLSMWRARS 105

RESULT 3
US-09-489-847-309
; Sequence 309, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 309
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-309

Query Match
Best Local Similarity 15.7%; Score 62.5; DB 4; Length 331;
Matches 16; Conservative 13; Mismatches 14; Indels 9; Gaps 2;

QY 21 KMPKLIIDTFFSIVAPALTAVLSQCLRCSLWLVGARLXPCGKVEGMDVWR 72
DB 53 ELPENILLEFTHV-PARQLLNCRIVCSLWR-----DLIDMLTWKRX 95

RESULT 4
US-09-489-039A-12478
; Sequence 12478, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
```

```
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12478
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12478

Query Match
Best Local Similarity 15.3%; Score 61; DB 4; Length 240;
Matches 23; Conservative 11; Mismatches 23; Indels 18; Gaps 4;

QY 13 EKAGKWC---KMPKLIIDTFFSIVAPA-----LTAVLSQCLRCSLWLV-----GAR 56
DB 143 QRAALARCLVRQPVLLIDFFSALDPAALQEMLSLVADVCQQQLTLLMVSHSVDAAR 202
QY 57 LXPCGKV--EGMDVW 69
DB 203 IAPRSMVVAEGRIWV 217

RESULT 5
US-09-172-952-14
; Sequence 14, Application US/09172952
; Patent No. 6368793
; GENERAL INFORMATION:
; APPLICANT: Hoch, James
; APPLICANT: Dartois, Veronique
; TITLE OF INVENTION: METABOLIC SELECTION METHODS
; FILE REFERENCE: 234/191
; CURRENT APPLICATION NUMBER: US/09/172,952
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 439
; TYPE: PRT
; ORGANISM: YiaX2
US-09-172-952-14

Query Match
Best Local Similarity 32.7%; Score 60.5; DB 4; Length 439;
Matches 18; Conservative 8; Mismatches 22; Indels 7; Gaps 2;

QY 7 MFSMFMEKAGKWCCKMPKLIIDTFFSIVAPALTAVLSQCLRCSLWLVGARLXPCG 61
DB 296 LFSSLSDRCK-----RKLFCVCLP--LIGFALCNFLSVALKNQIWLSYAALVCGG 343

RESULT 6
US-09-252-991A-25416
; Sequence 25416, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25416
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25416
```


APPLICANT: Gupta, Ashwani
APPLICANT: Vyas, Tejal
APPLICANT: McCallum, Kirk
APPLICANT: Fan, Emel
TITLE OF INVENTION: CLONED GLUCAGON-LIKE PEPTIDE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/845,546
FILING DATE: 24-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Friebe, Thomas E.
REGISTRATION NUMBER: 29,258
REFERENCE/DOCKET NUMBER: 8607-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-845-546-12

Query Match 14.1%; Score 56; DB 3; Length 553;
Best Local Similarity 30.9%; Pred. No. 41;
Matches 17; Conservative 8; Mismatches 22; Indels 8; Gaps 2;
QY 28 DTFEFTVAPALTAVALSCQLRCSLWLVG-----ARLXPCGKVEG--MDVWRRRWS 74
DB 23 ELPMGIPAPWGTSPLSFHRKCSLWAPGRPFLTLVLLVLSIKQVTGSLLETTRKWA 77

RESULT 12
US-09-907-794A-339
Sequence 339, Application US/09907794A
Patent No. 6635468
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Quang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kijavlin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James

APPLICANT: Paoni, Nicholas P.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
ACIDS
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/907,794A
CURRENT FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 339
LENGTH: 772
TYPE: PRT
ORGANISM: Homo Sapien
US-09-907-794A-339

Query Match 14.1%; Score 56; DB 4; Length 772;
Best Local Similarity 32.5%; Pred. No. 60;
Matches 13; Conservative 6; Mismatches 15; Indels 6; Gaps 1;
QY 32 SIVAPALTAVALSCQLRCSL-----WLVGARLXPCGKVEG 65
DB 8 ALLRPAFLTLGLSLGCSLLRLVRSWTQGEEDPCVEAVG 47

RESULT 13
US-09-905-125A-339
Sequence 339, Application US/09905125A
Patent No. 6664376
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen

APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/905/125A
CURRENT FILING DATE: 2001-07-12
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 339
LENGTH: 772
TYPE: PRT
ORGANISM: Homo Sapien
US-09-905-125A-339
Query Match 14.1%; Score 56; DB 4; Length 772;
Best Local Similarity 32.5%; Pred. No. 60;
Matches 13; Conservative 6; Mismatches 15; Indels 6; Gaps 1;
Qy 32 SIVAPALTAVLSQRLSL-----WLVGARLYPCGKVEG 65
Db 8 ALLRPALPLTLGLSLGSLRLVSWIQEGEDPCVEAVG 47

RESULT 14
US-09-902-775A-339
Sequence 339, Application US/09902775A
Patent No. 6686451
GENERAL INFORMATION:
APPLICANT: Genentech, Inc.
APPLICANT: Ashkenazi, Avi
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/902,775A
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 339
LENGTH: 772
TYPE: PRT

Search completed: March 14, 2004, 21:49:41
Job time : 10.6966 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 19:09:12 ; Search time 30.6207 Seconds
(without alignments)
682.823 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 398

Sequence: 1 ARAYXKMFMEKAGKWC.....ARLXPCQKVEGMDVRRWS 74

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : 1 Genesep29Jan04:*

2 Genesep1980s:*

3 Genesep1990s:*

4 Genesep2000s:*

5 Genesep2001s:*

6 Genesep2002s:*

7 Genesep2003as:*

8 Genesep2003bs:*

9 Genesep2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	394	99.0	74	3	AAB08845
2	65.5	16.5	142	6	ADA54405
3	65.5	16.5	253	3	AB52094
4	65.5	16.5	383	4	AB66060
5	65.5	16.5	383	6	ABO32714
6	64.5	16.2	177	4	AA24251
7	64.5	16.2	493	4	AB93101
8	64.5	16.2	493	5	AB97427
9	62.5	15.7	293	5	ABP43753
10	62.5	15.7	331	3	AA91629
11	61.5	15.5	205	5	ABR40427
12	61.5	15.5	205	5	ABR40504
13	61.5	15.5	243	5	ABR40570
14	61.5	15.5	268	4	AA79104
15	61.5	15.5	280	4	AB11361
16	61.5	15.5	280	4	AA80088
17	61.5	15.5	330	4	AA95481
18	61.5	15.5	387	3	AA805764
19	61.5	15.5	438	4	AA83081
20	61.5	15.5	460	4	AA83079
21	61.5	15.5	504	7	ADB65718
22	61.5	15.5	504	7	ADB65718
23	61.5	15.5	760	6	ABR41882
24	60.5	15.2	300	4	ABG21739
25	60.5	15.2	439	3	AA70726
26	60.5	15.2	439	5	AA50969

26	60.5	15.2	439	5	AAU77016	Aau77016 2,5-diket
27	60.5	15.2	439	5	AAO19977	Aao19977 Protein o
28	60	15.1	43	3	AB34861	Ab34861 Human sec
29	60	15.1	109	4	AAU43031	Aau43031 Propionib
30	60	15.1	109	6	ABM39550	Abm39550 Propionib
31	60	15.1	477	4	AB61772	Ab61772 Drosophil
32	60	15.1	477	4	AB61772	Ab61772 Drosophil
33	59.5	14.9	189	3	AA74817	Aay74817 Neisseria
34	59.5	14.9	190	3	AA74815	Aay74815 Neisseria
35	59.5	14.9	620	4	ABG15968	Abg15968 Novel hum
36	59.5	14.9	1211	4	AB11434	Ab11434 P. pastor
37	59	14.8	333	7	ABM73765	Abm73765 DNA clone
38	59	14.8	386	7	ABM74366	Abm74366 DNA clone
39	58.5	14.7	60	3	AA27563	Aa27563 Human sec
40	58.5	14.7	141	4	AAO09875	Aao09875 Human pol
41	58.5	14.7	224	5	ABP69507	Abp69507 Human pol
42	58.5	14.7	224	5	ABU61007	Abu61007 Lung spec
43	58.5	14.7	224	6	ABP97741	Abp97741 Amino aci
44	58.5	14.7	255	6	ABR43255	Ab43255 Human PMW
45	58.5	14.7	280	3	AA42989	Aab42989 Human ORF

ALIGNMENTS

RESULT 1
AAB08845
ID AAB08845 standard; protein; 74 AA.
XX
AC AAB08845;
XX
DT 02-JAN-2001 (first entry)
XX
DE A human MDM2 interacting polypeptide (MDMIP).
XX

Human; MDM2 interacting polypeptide; MDMIP; MDM2; cell cycle progression;
cell differentiation; cancer; sarcoma; glioma; squamous cell carcinoma;
breast cancer; astrocytoma; leukemia; lymphoma; tumorigenesis;
gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 5 /note= "unknown amino acid encoded by NAA"

FT Misc-difference 58 /note= "unknown amino acid encoded by TNT"

FT Misc-difference 58 /note= "unknown amino acid encoded by TNT"

XX WO200050590-A1.

XX 31-AUG-2000.

XX 23-FEB-2000; 2000WO-US004582.

XX 23-FEB-1999; 99US-0121192P.

XX 03-MAR-1999; 99US-0122643P.

XX 22-FEB-2000; 2000US-00510252.

XX (CURA-) CURAGEN CORP.

XX Nandabalan K, Yang M, Schulz VP;

XX WPI; 2000-558398/51.

XX N-PSDB; AAA75041.

XX Novel MDM2 interacting protein useful for treating or preventing
PT disorders involving aberrant levels of MDM2 and/or MDM-interacting
PT proteins, comprises a specific amino acid sequence.

XX Claim 10; Fig 1; 78pp; English.

XX The present sequence represents a human MDM2 interacting polypeptide
CC (MDMIP). MDMIP was identified using a yeast two hybrid system, using a

CC fragment of MD2 as the bait protein. The MD2IP polypeptide is useful for
 CC detecting and removing MD2 polypeptides in a biological sample by
 CC forming MD2-MD2IP complexes. MD2IP and MD2 are useful to identify
 CC compounds or other agents which modulate the activity of MD2 and/or
 CC MD2IP-mediated processes. Agents that modulate the function of MD2IP/MD2
 CC complexes are useful for treating and preventing a disease or disorder
 CC involving aberrant levels of MD2 or MD2IP. MD2IP is also useful for
 CC treating diseases caused by aberrant levels of expression of MD2 genes,
 CC such as disorders of cell cycle progression, cell differentiation, and
 CC transcriptional control, including cancers such as human sarcoma, glioma,
 CC squamous cell carcinoma, breast cancer, astrocytoma, leukemia and
 CC lymphoma, and tumorigenesis. MD2IP and MD2 nucleic acids are useful in
 CC gene therapy

XX Sequence 74 AA;

Query Match 99.0%; Score 394; DB 3; Length 74;
 Best Local Similarity 100.0%; Pred. No. 1.6e-42;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARAYXQMFNFMEKAGCKWCKMPKLIIDTFFSIVAPALTAVLSQRLSMLVGLARLXPC 60
 DB 1 ARAYXQMFNFMEKAGCKWCKMPKLIIDTFFSIVAPALTAVLSQRLSMLVGLARLXPC 60

QY 61 GKVEGMDVWRRWS 74
 DB 61 GKVEGMDVWRRWS 74

RESULT 2
 ADA54405
 ID ADA54405 standard; protein; 142 AA.
 AC ADA54405;
 DT 20-NOV-2003 (first entry)
 DE Human protein, SEQ ID 1973.

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
 KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
 KW inflammatory disease; osteoporosis; neurological disease.

OS Homo sapiens.
 XX EP1293569-A2.
 XX 19-MAR-2003.
 XX 21-MAR-2002; 2002EP-00006586.
 XX 14-SEP-2001; 2001JP-00328381.
 XX 24-JAN-2002; 2002US-0350435P.
 XX (HELI-) HELIX RES INST.
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;
 XX WPI; 2003-395539/38.
 DR N-PSDB; ADA52766.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory
 PT and/or membrane proteins, useful for developing medicines for diseases in
 PT which the gene is involved, or as target molecules for gene therapy.

PS Claim 14; SEQ ID NO 1973; 205pp; English.

XX The present invention relates to novel human secretory or membrane
 CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
 CC ADA54071). The coding sequences are useful in the gene therapy of

CC diseases caused by abnormalities of the proteins, e.g. cancer,
 CC inflammatory diseases, osteoporosis or neurological disease.

SQ Sequence 142 AA;

Query Match 16.5%; Score 65.5; DB 6; Length 142;
 Best Local Similarity 30.3%; Pred. No. 3.2;
 Matches 20; Conservative 9; Mismatches 22; Indels 15; Gaps 4;

QY 19 WCKMPKLIIDTPESI--VAPAL-----TAVLSQRLSMLVGLARLXPCGKVEGMDV-- 68
 DB 2 WALNPRFVTPFLKLCNISPGLPGRGTSVLP---PCCLWIRALRPATWLCLEGMEDV 58

QY 69 --WRRR 72
 DB 59 GQWRRR 64

RESULT 3
 AAB52094
 ID AAB52094 standard; protein; 253 AA.
 AC AAB52094;

DT 21-FEB-2001 (first entry)

DE Gene 44 human secreted protein homologous amino acid sequence #143.

XX Human; secreted protein; cytostatic; immunosuppressive; nootropic;
 KW neuroprotective; antiviral; antiallergic; hepatotropic; antidiabetic;
 KW antiinflammatory; antitumor; vulnerary; anticonvulsant; antibacterial;
 KW antifungal; antiparasitic; cardiant; cancer; immune disease; allergy;
 KW cardiovascular disorder; wound healing; infection; neurological disease.

XX Caenorhabditis elegans.

OS WO2000061596-A1.

XX 19-OCT-2000.

XX 06-APR-2000; 2000WO-US008983.

XX 09-APR-1999; 99US-0128703P.

XX 14-JAN-2000; 2000US-0176068P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX Ruben SM, Komatsoulis G;

XX WPI; 2000-611865/58.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful in
 PT the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases.

XX Disclosure; Page 75-76; 505pp; English.

XX Polynucleotide sequences AAC95521 - AAC95570 represent cDNA encoding 50
 CC human secreted proteins AAB52012 - AAB52161. Sequences AAB52062 -
 CC AAB52103 represent alternative polypeptides encoded by the genes, and
 CC amino acid sequences with which they share homology. The genes and
 CC proteins have activities dependent on the tissues and cells in which they
 CC are expressed. Examples of their activities include cytostatic;
 CC immunosuppressive; nootropic; neuroprotective; antiviral; antiallergic;
 CC hepatotropic; antidiabetic; antiinflammatory; antitumor; vulnerary;
 CC anticonvulsant; antibacterial; antifungal; antiparasitic; and cardiant.
 CC The secreted proteins, polypeptides, antagonists and agonists may be
 CC useful in treating, preventing and/or diagnosing diseases and disorders
 CC such as cancer, particularly breast and ovarian cancer, and other cancers
 CC of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract,
 CC liver, lung, or urogenital. Immune disorders such as Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,

CC The invention relates to secreted polypeptide-related proteins and
CC nucleic acids (RANGO and INTERCEPT proteins and nucleic acids). The
CC nucleic acids, proteins and antibodies specific to the proteins are
CC useful in screening assays, predictive medicine (e.g. diagnostic assays,
CC prognostic assays, monitoring clinical trials and pharmacogenetics) and
CC prophylactic and therapeutic methods. The sequences are used in
CC diagnosing, preventing or treating proliferative disorders (e.g.
CC cancers), hormonal disorders (e.g. diabetes or thyroid disorders), immune
CC disorders (e.g. multiple sclerosis or lupus), neurological disorders
CC (e.g. Alzheimer's disease or Parkinson's disease), cardiovascular
CC disorders (e.g. myocardial infarction or congestive heart disease), blood
CC platelet disorders (e.g. thrombocytopenia or anaemia) and disorders
CC involving blood vessels (e.g. atherosclerosis or vasculitis). The nucleic
CC acids may also be used in chromosome mapping, tissue typing and forensic
CC biology, and as surrogate markers. This sequence represents a secreted
CC polypeptide-related protein of the invention. Note: The sequence data for
CC this patent was obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 383 AA;

Query Match 16.5%; Score 65.5; DB 6; Length 383;
Best Local Similarity 30.2%; Pred. No. 9.7;
Matches 13; Conservative 8; Mismatches 19; Indels 3; Gaps 1;
QY 18 CWCXMPKLIIDTPFSIVAPALTAIVLSQRLSCRLSLWLVGRLXPC 60
DB 185 CWCLSVLMFLPHNAYKILATIGISCLIACLVYLL---LSPC 224

RESULT 6
AAW24251
ID AAW24251 standard; protein; 177 AA.
XX
AC AAW24251;
XX
DT 12-OCT-2001 (first entry)
XX
DE Human EST encoded protein SEQ ID NO: 1776.
XX
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW diagnostics; forensic test; gene mapping; genetic disorder; biodiversity;
KW gene therapy; nutrition.
XX
OS Homo sapiens.
XX
PI WO200154477-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US002687.
XX
PR 25-JAN-2000; 2000US-00491404.
PR 17-JUL-2000; 2000US-00617746.
PR 03-AUG-2000; 2000US-00631451.
PR 15-SEP-2000; 2000US-00663870.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;
PI Cao Y, Drmanac RA, Zhang J, Werhman T;
XX
DR WPI; 2001-476164/51.
DR N-PSDB; AAH98910.
XX
PT Isolated polypeptide for treatment of diseases, diagnostics, raising
PT antibodies and research use.
XX
PS Claim 20; Page 1168; 1275pp; English.
XX
CC The present invention provides the protein and coding sequences of novel
CC proteins from a variety of organisms, including human, dog, cat, horse,

CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC from the organism of interest. They can be used in diagnostics,
CC forensics, gene mapping, identification of mutations, to assess
CC biodiversity and for nutritional purposes. The present sequence is a
CC protein of the invention
XX
SQ Sequence 177 AA;

Query Match 16.2%; Score 64.5; DB 4; Length 177;
Best Local Similarity 28.3%; Pred. No. 5.5;
Matches 17; Conservative 10; Mismatches 24; Indels 9; Gaps 3;
QY 18 CWCXMPKLIIDTPFSIVAPALTAIVLSQRLSCRLSLWLVGRLXPCGVGMDVWRRW 73
DB 4 CWC-----VLGTPGVAMVLLHTTIFPCVAFRSQLLTLCSTLLSLRLQGVVEVKRW 58

RESULT 7
AAB93101
ID AAB93101 standard; protein; 493 AA.
XX
AC AAB93101;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:11958.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PI EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-00116126.
XX
PR 29-JUL-1999; 99JP-00248036.
PR 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
PR 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
PT cDNAs.
XX
PS Claim 8; SEQ ID NO 11958; 2537pp + Sequence Listing; English.
XX
CC The present invention describes primer sets for synthesising 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the

Best Local Similarity 28.3%; Pred. No. 17;
Matches 17; Conservative 10; Mismatches 24; Indels 9; Gaps 3;

Db 15 ELPENILLEFTHV-PARQLLLNCRLVCSLWR-----DLIDLMTLWKRK 57

CC multiple sclerosis, respiratory disorders (e.g. asthma), or infectious
CC diseases caused by bacterial, parasitic, viral or fungal agents. The
CC nucleic acids are also useful for chromosome identification, radiation
CC hybrid mapping or long-range restriction mapping. The polypeptide,
CC or preservative to increase or decrease storage capabilities, fat content
CC or other nutritional components. The sequences given in records ABR40409-
CC ABR40590 and AB282464-AB282611 represent human secreted proteins and the
CC genes encoding them
XX
SQ Sequence 205 AA;
Query Match 15.5%; Score 61.5; DB 5; Length 205;
Best Local Similarity 28.2%; Pred. No. 16;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;
QY 4 YXKMFSEFM--EKAGCKWCKMPKLIITPFSIVAPALTAVLSQLRCSLWVGARLXP-- 59
DB 9 YSKNISLMMNFQPPSKAW-RASQMTTFILFPFSFTGVL-CTLAITIW-----RLKPSA 62
QY 60 -CGKVEGMDVW 69
DB 63 DCGPFRGLPLF 73
RESULT 12
ID ABR40504 standard; protein; 205 AA.
XX
AC ABR40504;
XX
DT 13-JUN-2003 (first entry)
XX
DE Human secreted protein #SEQ ID 254.
XX
KW Human; secreted protein; anti-HIV; nontropic; neuroprotective;
KW antiangiinal; immunosuppressive; immunomodulator; cytostatic; cardiant;
KW hepatotropic; antiinflammatory; antiallergic; antidiabetic;
KW gastrointestinal; antiinfertility; nephrotropic; virucide; hypotensive;
KW vasotropic; dermatological; osteopathic; antiarthritic; antiparkinsonian;
KW antiasthmatic; antipsoriatic; cerebroprotective; antibacterial;
KW fungicide; antiparasitic; protein therapy; gene therapy; cancer;
KW hyperproliferative disorder; leukaemia; autoimmune disorder;
KW immunodeficiency; HIV; infertility; uterine disorders; endometriosis;
KW amenorrhea; ocular disorder; neurological disorder; wound healing;
KW Huntington's disease; gastrointestinal disorder; inflammatory disease;
KW Crohn's disease; kidney disorder; renal failure; diabetes; obesity;
KW cirrhosis; hepatitis; cardiovascular disorder; angina; osteoarthritis;
KW rheumatoid arthritis; psoriasis; multiple sclerosis; asthma;
KW respiratory disorder; infectious disease; chromosome identification;
KW food additive; nutrition.
XX
OS Homo sapiens.
XX
PN WO200268628-A1.
XX
PD 06-SEP-2002.
XX
PF 21-FEB-2002; 2002WO-US005301.
XX
PR 23-FEB-2001; 2001US-0270625P.
PR 12-JUL-2001; 2001US-030417P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Komatsoulis G, Baker KP, Fiscella M, Moore PA, Wei P;
PI Duan DR, Shi Y, Gupta R;
XX
DR WPI; 2002-750417/81.
DR N-PSDB; AB282559.
XX
PT New human secreted proteins and nucleic acids, useful for preventing,
PT treating, or ameliorating a medical condition, e.g. cancer, autoimmune

PT disorder, HIV, Parkinson's disease, Alzheimer's disease, diabetes,
PT obesity or cirrhosis.
XX
PS Claim 11; Page 812-813; 873pp; English.
XX
CC The invention relates to novel human secreted proteins and the genes
CC encoding them. Genes and proteins of the invention may be useful for
CC preventing, treating or ameliorating medical conditions e.g. by protein
CC or gene therapy. These conditions include cancer and hyperproliferative
CC disorders, immune cell proliferative disorders (e.g. leukaemia),
CC autoimmune disorders, immunodeficiencies (e.g. HIV infection),
CC infertility, placental and uterine disorders (e.g. endometriosis),
CC amenorrhea, ocular disorders, neurological disorders (e.g. Huntington's
CC disease), wound healing, gastrointestinal system disorders, particularly
CC inflammatory diseases (e.g. Crohn's disease), kidney disorders (e.g.
CC renal failure), diabetes, obesity, cirrhosis, hepatitis, cardiovascular
CC disorders (e.g. angina), rheumatoid arthritis, osteoarthritis, psoriasis,
CC multiple sclerosis, respiratory disorders (e.g. asthma), or infectious
CC diseases caused by bacterial, parasitic, viral or fungal agents. The
CC nucleic acids are also useful for chromosome identification, radiation
CC hybrid mapping or long-range restriction mapping. The polypeptide,
CC or preservative to increase or decrease storage capabilities, fat content
CC or other nutritional components. The sequences given in records ABR40409-
CC ABR40590 and AB282464-AB282611 represent human secreted proteins and the
CC genes encoding them
XX
SQ Sequence 205 AA;
Query Match 15.5%; Score 61.5; DB 5; Length 205;
Best Local Similarity 28.2%; Pred. No. 16;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;
QY 4 YXKMFSEFM--EKAGCKWCKMPKLIITPFSIVAPALTAVLSQLRCSLWVGARLXP-- 59
DB 9 YSKNISLMMNFQPPSKAW-RASQMTTFILFPFSFTGVL-CTLAITIW-----RLKPSA 62
QY 60 -CGKVEGMDVW 69
DB 63 DCGPFRGLPLF 73
RESULT 13
ID ABR40570 standard; protein; 243 AA.
XX
AC ABR40570;
XX
DT 13-JUN-2003 (first entry)
XX
DE Human secreted protein #SEQ ID 320.
XX
KW Human; secreted protein; anti-HIV; nontropic; neuroprotective;
KW antiangiinal; immunosuppressive; immunomodulator; cytostatic; cardiant;
KW hepatotropic; antiinflammatory; antiallergic; antidiabetic;
KW gastrointestinal; antiinfertility; nephrotropic; virucide; hypotensive;
KW vasotropic; dermatological; osteopathic; antiarthritic; antiparkinsonian;
KW antiasthmatic; antipsoriatic; cerebroprotective; antibacterial;
KW fungicide; antiparasitic; protein therapy; gene therapy; cancer;
KW hyperproliferative disorder; leukaemia; autoimmune disorder;
KW immunodeficiency; HIV; infertility; uterine disorders; endometriosis;
KW amenorrhea; ocular disorder; neurological disorder; wound healing;
KW Huntington's disease; gastrointestinal disorder; inflammatory disease;
KW Crohn's disease; kidney disorder; renal failure; diabetes; obesity;
KW cirrhosis; hepatitis; cardiovascular disorder; angina; osteoarthritis;
KW rheumatoid arthritis; psoriasis; multiple sclerosis; asthma;
KW respiratory disorder; infectious disease; chromosome identification;
KW food additive; nutrition.
XX
OS Homo sapiens.
XX
PN WO200268628-A1.
XX


```

PD 06-SEP-2002.
XX
XX 21-FEB-2002; 2002WO-US005301.
XX
XX 23-FEB-2001; 2001US-0270625P.
PR 12-JUL-2001; 2001US-0304417P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Komatsoulis G, Baker KP, Fiscella M, Moore PA, Wei P;
PI Duan DR, Shi Y, Gupta R;
XX
XX WPI; 2002-750417/81.
XX
XX New human secreted proteins and nucleic acids, useful for preventing,
PT treating, or ameliorating a medical condition, e.g. cancer, autoimmune
PT disorder, HIV, Parkinson's disease, Alzheimer's disease, diabetes,
PT obesity or cirrhosis.
XX
XX Claim 11; Page 855; 873pp; English.
XX
XX The invention relates to novel human secreted proteins and the genes
XX encoding them. Genes and proteins of the invention may be useful for
XX preventing, treating or ameliorating medical conditions e.g. by protein
XX or gene therapy. These conditions include cancer and hyperproliferative
XX disorders, immune cell proliferative disorders (e.g. leukaemia),
XX autoimmune disorders, immunodeficiencies (e.g. HIV infection),
XX infertility, placental and uterine disorders (e.g. endometriosis),
XX amenorrhea, ocular disorders, neurological disorders (e.g. Huntington's
XX disease), wound healing, gastrointestinal system disorders, particularly
XX inflammatory diseases (e.g. Crohn's disease), kidney disorders (e.g.
XX renal failure), diabetes, obesity, cirrhosis, hepatitis, cardiovascular
XX disorders (e.g. angina), rheumatoid arthritis, osteoarthritis, psoriasis,
XX multiple sclerosis, respiratory disorders (e.g. asthma) or infectious
XX diseases caused by bacterial, parasitic, viral or fungal agents. The
XX nucleic acids are also useful for chromosome identification, radiation
XX hybrid mapping or long-range restriction mapping. The polypeptide,
XX polynucleotide, agonist or antagonist may also be used as a food additive
XX or preservative to increase or decrease storage capabilities, fat content
XX or other nutritional components. The sequences given in records ABR40409-
XX ABR40590 and ABR28464-ABR28611 represent human secreted proteins and the
XX genes encoding them
XX
XX Sequence 243 AA;
SQ
Query Match 15.5%; Score 61.5; DB 5; Length 243;
Best Local Similarity 28.2%; Pred. No. 19;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;
QY 4 YXKMFNFM--EKAGKWCCKMPKLIIDTPFSIVAPALTAVLSQLCRCSLWVGARLXP-- 59
Db 86 YSKNISLNMNFQPPSKAW-RASQMTFFIFLLFPFSFTGVL-CTLAITW----RLKPSA 139
QY 60 -CGKVEGMDVW 69
Db 140 DCGPFRGLPLF 150
RESULT 14
AAM79104
ID AAM79104 standard; protein; 268 AA.
XX
XX AAM79104;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human protein SEQ ID NO 1766.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorder; arthritis; inflammation.
XX

```

```

OS Homo sapiens.
XX
XX WO200157190-A2.
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX
XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
DR N-PSDB; AAK52237.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX in diagnosis and gene therapy.
XX
XX Claim 20; Page 4113-4114; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX sequence listing were missing at the time of publication
XX
XX Sequence 268 AA;
SQ
Query Match 15.5%; Score 61.5; DB 4; Length 268;
Best Local Similarity 28.2%; Pred. No. 21;
Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;
QY 4 YXKMFNFM--EKAGKWCCKMPKLIIDTPFSIVAPALTAVLSQLCRCSLWVGARLXP-- 59
Db 72 YSKNISLNMNFQPPSKAW-RASQMTFFIFLLFPFSFTGVL-CTLAITW----RLKPSA 125
QY 60 -CGKVEGMDVW 69
Db 126 DCGPFRGLPLF 136
RESULT 15
ABBI1361
ID ABBI1361 standard; peptide; 280 AA.
XX
XX ABBI1361;
AC
XX
XX 11-JAN-2002 (first entry)
DT
XX
XX Human LAK-4p homologue, SEQ ID NO:1731.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
XX haematopoiesis regulation; tissue growth; immunomodulator; activin;
XX inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
XX

```

KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytoskeletal; osteopathic; vasotropic; cardiant; virucide; antibacterial;
 KW antifungal; vulnery; antiulcer.
 XX Homo sapiens.
 XX WO200157188-A2.
 XX 09-AUG-2001.
 XX 05-FEB-2001; 2001WO-US003800.
 XX 03-FEB-2000; 2000US-00496914.
 XX 27-APR-2000; 2000US-00560875.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-457740/49.
 XX N-PSDB; ABA08605.
 XX Human proteins and DNA encoding sequences useful for preventing, treating
 XX or ameliorating a medical condition in a mammalian subject e.g. arthritis
 XX and cancer.
 XX Claim 20; Page 173; 1963pp; English.
 XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 XX sequences ABA0825-ABA09574 represent nucleic acids encoding them. The
 XX invention also relates to vectors and recombinant host cells comprising a
 XX nucleotide of the invention, methods of producing the novel polypeptides,
 XX antibodies against the polypeptides, methods of detecting the nucleotides
 XX or polypeptides in a sample, and methods of identifying compounds which
 XX bind to polypeptides of the invention. Although novel, many of the
 XX polypeptides of the invention have homology to known proteins, thereby
 XX giving an insight into their probable biological activities, and hence
 XX potential therapeutic applications. The polypeptides of the invention may
 XX have various activities, including cytokine, cell proliferation or cell
 XX differentiation activities; stem cell growth factor activity;
 XX haematopoiesis regulatory activity; tissue growth activity;
 XX immunomodulatory activity; activin- or inhibin-related activities;
 XX chemotactic or chemokinetic activities; haemostatic, thrombotic or
 XX thrombolytic activities; receptor or ligand activities; or may be
 XX involved in oncogenesis; cancer cell proliferation or metastasis.
 XX Depending on their biological activities, polypeptides and nucleotides of
 XX the invention are useful for preventing, treating or ameliorating medical
 XX conditions, e.g., by protein or gene therapy. Such conditions include
 XX cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 XX disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 XX proliferative retinopathy, atherosclerosis, coronary heart disease,
 XX arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
 XX vascular growth. Polypeptides involved with tissue regeneration and
 XX repair (or nucleic acids encoding them) may be used to promote wound
 XX healing (e.g., of burns, incisions and ulcers), while those with
 XX immunomodulatory activities may be used in the treatment of viral,
 XX bacterial and fungal infections in addition to immune disorders.
 XX Polypeptides with growth factor activity may be used in cell cultures to
 XX promote cell growth. For example, such polypeptides may be used to
 XX manipulate stem cells in culture to give rise to neuroepithelial cells
 XX that can be used to augment or replace cells damaged by illness,
 XX autoimmune disease or accidental damage. The polypeptides and nucleotides
 XX may also be used in the diagnosis of the above conditions, and in drug
 XX screening techniques. The present sequence represents a novel human
 XX polypeptide of the invention

SQ Sequence 280 AA;

Query Match 15.5%; Score 61.5; DB 4; Length 280;
 Best Local Similarity 28.2%; Pred. No. 22;
 Matches 20; Conservative 13; Mismatches 27; Indels 11; Gaps 5;

QY 4 YXMFSPFM--EKAGCKCKPKLIIDTFFSIVAPALTAVLSQRCISLWLVGARLXP-- 59

DB 84 YSKNISLMMNFOPPSKAM-RASQMTFFIFLFFPSFTGVL-CTLAITW-----RLKPSA 137

QY 60 -CGKVEGMDVW 69

DB 138 DCGPFRGLPLF 148

Search completed: March 14, 2004, 21:44:53
 Job time : 35.6207 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:48:49 ; Search time 14.0345 Seconds
(without alignments)
1663.641 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 74

Sequence: 1 ARAYXKMFMEKAGKWC.....ARLXPCGVGMGVRRWS 74

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL 25.*

- 1: sp_archae.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organalle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archae.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	10.8	1854	11 Q63131	Q63131 rattus norv
2	8	10.8	2317	11 Q63130	Q63130 rattus norv
3	8	10.8	2338	11 Q63132	Q63132 rattus norv
4	7	9.5	98	16 Q92TK8	Q92TK8 rhizobium m
5	7	9.5	124	16 Q7V2A6	Q7V2A6 prochloroco
6	7	9.5	144	12 Q91UB3	Q91UB3 human polio
7	7	9.5	200	16 Q82E32	Q82E32 streptomyc
8	7	9.5	203	16 Q8K450	Q8K450 streptomyc
9	7	9.5	217	16 Q7WHY5	Q7WHY5 bordetella
10	7	9.5	217	16 Q7WH915	Q7WH915 bordetella
11	7	9.5	228	16 Q82UM7	Q82UM7 nitrosomona
12	7	9.5	261	16 Q92EW9	Q92EW9 listeria in
13	7	9.5	261	16 Q8YA48	Q8YA48 listeria mo
14	7	9.5	281	16 Q8EC47	Q8EC47 shewanella
15	7	9.5	303	12 Q91LW8	Q91LW8 human coxa
16	7	9.5	312	16 Q50525	Q50525 streptomyc

17	7	9.5	325	16 Q89NK1	Q89NK1 bradyrhizob
18	7	9.5	377	16 Q88WH4	Q88WH4 lactobacill
19	7	9.5	398	16 Q7WEG4	Q7WEG4 bordetella
20	7	9.5	398	16 Q7W347	Q7W347 bordetella
21	7	9.5	407	16 Q89X37	Q89X37 bradyrhizob
22	7	9.5	434	16 Q8VKK1	Q8VKK1 mycobacteri
23	7	9.5	434	16 Q7U1V1	Q7U1V1 mycobacteri
24	7	9.5	444	5 Q20584	Q20584 caenorhabdi
25	7	9.5	464	16 Q81W2	Q81W2 bacillus an
26	7	9.5	464	16 Q81A66	Q81A66 bacillus ce
27	7	9.5	537	16 Q8YCM4	Q8YCM4 brucella me
28	7	9.5	537	16 Q8FV2	Q8FV2 brucella su
29	7	9.5	568	16 Q8NSC0	Q8NSC0 corynebacte
30	7	9.5	669	5 Q9V5Z3	Q9V5Z3 drosophila
31	7	9.5	923	2 Q8KHD3	Q8KHD3 nocardia ae
32	7	9.5	1318	5 Q9NE11	Q9NE11 leishmania
33	7	9.5	1326	16 Q8DM52	Q8DM52 synchococc
34	7	9.5	2207	12 Q7T7N8	Q7T7N8 human coxa
35	7	9.5	6797	2 Q9X993	Q9X993 streptomyc
36	6	8.1	20	10 Q9S885	Q9S885 lupinus alb
37	6	8.1	26	15 Q9PXA6	Q9PXA6 human immun
38	6	8.1	29	12 Q9WMY2	Q9WMY2 human echov
39	6	8.1	29	12 Q9WMY3	Q9WMY3 human echov
40	6	8.1	29	12 Q9W8F9	Q9W8F9 human echov
41	6	8.1	29	12 Q9WMY4	Q9WMY4 human echov
42	6	8.1	30	12 Q9WMY0	Q9WMY0 human echov
43	6	8.1	31	12 Q9WMY6	Q9WMY6 human echov
44	6	8.1	31	12 Q9WMX6	Q9WMX6 human echov
45	6	8.1	33	12 Q9WMX7	Q9WMX7 human echov

ALIGNMENTS

RESULT 1

Q63131 PRELIMINARY; PRT; 1854 AA.

AC Q63131
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
 DE Rat lung-derived L01 c-ras-1 proto-oncogene.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90219211; PubMed=2139140;
 RA Matsushime H., Shibuya M.;
 FT "tissue-specific expression of rat c-ras-1 gene and partial structural
 RT similarity of its predicted products with sev protein of Drosophila
 RT melanogaster.";
 RL J. Virol. 64:2117-2125 (1990).
 DR EMBL; M35105; AAA40967.1; .
 DR GO; GO:0016020; C:membrane; IEA.
 DR InterPro; IPR0031961; FN.III.
 DR InterPro; IPR008957; FN.III-like.
 DR InterPro; IPR000033; Ldl_receptor_rep.
 DR Pfam; PF00041; fn3; 7.
 DR SMART; SMO0060; FN3; 8.
 DR SMART; SMO0135; LY; 2.
 SQ SEQUENCE 1854 AA; 207623 MW; CEDACB033F41BB8A CRC64;

Query Match 10.8%; Score 8; DB 11; Length 1854;

Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 29 TPFSIVAP 36

Db 413 TPFSIVAP 420

```

RESULT 2
Q63130 Q63130 PRELIMINARY; PRT; 2317 AA.
AC Q63130;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rat lung-derived c-ros-1 proto-oncogene.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90219211; PubMed=2139140;
RA Matsushime H., Shibuya M.;
RT "Issue-specific expression of rat c-ros-1 gene and partial structural
RT similarity of its predicted products with sev protein of Drosophila
RT melanogaster.";
RL J. Virol. 64:2117-2125(1990).
DR EMBL; M35104; AAA40966.1; -.
DR HSP; P11362; 1FGK.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000033; LdI_receptor_rep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00041; fn3; 7.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 8.
DR SMART; SM00135; LY; 2.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 2317 AA; 259095 MW; 4117901E3B0A6449 CRC64;

Query Match 10.8%; Score 8; DB 11; Length 2317;
Best Local Similarity 100.0%; Pred.No.20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 TPFSIVAP 36
DB 413 TPFSIVAP 420
|||||

RESULT 3
Q63132 Q63132 PRELIMINARY; PRT; 2338 AA.
AC Q63132;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rat heart-derived c-ros-1 proto-oncogene.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90219211; PubMed=2139140;
RA Matsushime H., Shibuya M.;
RT "Issue-specific expression of rat c-ros-1 gene and partial structural
RT similarity of its predicted products with sev protein of Drosophila
RT melanogaster.";
RL J. Virol. 64:2117-2125(1990).
DR EMBL; M35104; AAA40966.1; -.
DR HSP; P11362; 1FGK.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000033; LdI_receptor_rep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00041; fn3; 7.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 8.
DR SMART; SM00135; LY; 2.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 2317 AA; 259095 MW; 4117901E3B0A6449 CRC64;

Query Match 10.8%; Score 8; DB 11; Length 2317;
Best Local Similarity 100.0%; Pred.No.20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 TPFSIVAP 36
DB 413 TPFSIVAP 420
|||||

RESULT 3
Q63132 Q63132 PRELIMINARY; PRT; 2338 AA.
AC Q63132;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rat heart-derived c-ros-1 proto-oncogene.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90219211; PubMed=2139140;
RA Matsushime H., Shibuya M.;
RT "Issue-specific expression of rat c-ros-1 gene and partial structural
RT similarity of its predicted products with sev protein of Drosophila
RT melanogaster.";
RL J. Virol. 64:2117-2125(1990).
DR EMBL; M35106; AAA40968.1; -.
DR PIR; I73957; I73957.
DR HSP; P11362; 1FGK.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR000033; LdI_receptor_rep.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00041; fn3; 7.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00060; FN3; 8.
DR SMART; SM00135; LY; 2.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
KW ATP-binding; Kinase; Transferase; Tyrosine-protein kinase.
SQ SEQUENCE 2338 AA; 261653 MW; FF45FE561A058453 CRC64;

Query Match 10.8%; Score 8; DB 11; Length 2338;
Best Local Similarity 100.0%; Pred.No.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 TPFSIVAP 36
DB 413 TPFSIVAP 420
|||||

RESULT 4
Q92TK8 Q92TK8 PRELIMINARY; PRT; 98 AA.
ID Q92TK8
AC Q92TK8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein RB1503.
GN RB1503 OR SM520591.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=1021;
RC MEDLINE=21396508; PubMed=11481431;
RA Finan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
RA Golding B., Puehler A.;
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-
RT fixing endosymbiont Sinorhizobium meliloti.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894 (2001).
DR EMBL; AL603647; CAC49903.1; -.
DR PIR; G96029; G96029.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 98 AA; 11180 MW; C82C7B579CD36983 CRC64;

Query Match 9.5%; Score 7; DB 16; Length 98;
Best Local Similarity 100.0%; Pred.No.16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 LTAVLSC 44
|||||

```

```
DR Pfam; PF00073; xhv; 1.
FT NON_TER 1
PT NON_TER 144 144
SQ SEQUENCE 144 AA; 15956 MW; 783BAB2762A7430F CRC64;

Query Match 9.5%; Score 7; DB 12; Length 144;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 APALTAV 41
DB 58 APALTAV 64
|||||

RESULT 7
Q82E32 PRELIMINARY; PRT; 200 AA.
AC Q82E32;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN SAV4784.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis."
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AF05040; BAC72496.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 200 AA; 21837 MW; 8171FF813A96276C CRC64;

Query Match 9.5%; Score 7; DB 16; Length 200;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
DB 54 PALTAVL 60
|||||

RESULT 8
Q9K4B0 PRELIMINARY; PRT; 203 AA.
AC Q9K4B0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative tetr-family transcriptional regulator.
GN SCO5209 OR SC7E4.06.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;

DB 18 LTAVLSC 24

RESULT 5
Q7V2A6 PRELIMINARY; PRT; 124 AA.
AC Q7V2A6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN PM00576.
OS Prochlorococcus marinus subsp. pastoris (strain CCMP 1378 / MED4).
OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;
OC Prochlorococcus.
OX NCBI_TaxID=59919;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22825698; PubMed=12917642;
RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,
RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,
RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,
RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
RA Webb E.A., Zinser E.R., Chisholm S.W.;
RT "Genome divergence in two Prochlorococcus ecotypes reflects oceanic
RT niche differentiation."
RL Nature 424:1042-1047(2003).
DR EMBL; BX572091; CAE19035.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 124 AA; 14168 MW; D77C5955251EC137 CRC64;

Query Match 9.5%; Score 7; DB 16; Length 124;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 ALTAVALS 43
DB 92 ALTAVALS 98
|||||

RESULT 6
Q91UB3 PRELIMINARY; PRT; 144 AA.
AC Q91UB3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE VP3-VP1 capsid protein (Fragment).
GN Human poliovirus 3.
OS Human poliovirus 3.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12086;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S25-01;
RA Yoshida H., Matsuura K.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S25-01;
RX MEDLINE=20509843; PubMed=11055968;
RA Matsuura K., Ishikura M., Yoshida H., Nakayama T., Hasegawa S.,
RA Ando S., Horie H., Miyamura T., Kitamura T.;
RT "Assessment of poliovirus eradication in Japan: genomic analysis of
RT polioviruses isolated from river water and sewage in toyama
RT prefecture."
RL Appl. Environ. Microbiol. 66:5087-5091(2000).
DR EMBL; AB061302; BAB47384.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR001676; Rv.
DR InterPro; IPR008975; Viral_cap_coat.
```

```

RN SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RA Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RA "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2) / M145;
RX MEDLINE=97000351; PubMed=8943436;
RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RA "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
[4]
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2) / M145;
RX MEDLINE=21996410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.D.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders A., Sharp S., Squares R., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RA "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2).";
RL Nature 417:141-147(2002).
CC -1- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
DR EMBL; AL939122; CAB94594.1; -.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001647; HTH_Tetr.
DR Pfam; PF00440; tetr; 1.
DR PRINTS; PR00455; HTHVETR.
KW DNA-binding; Transcription regulation; Complete proteome.
SQ SEQUENCE 203 AA; 21209 MW; B924F30209AA0C31 CRC64;

Query Match 9.5%; Score 7; DB 16; Length 203;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 APALTAV 41
DB 111 APALTAV 117
|||||
PRT; 217 AA.

RESULT 9
Q7WH75 PRELIMINARY;
ID Q7WH75 AC Q7WH75
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN BB3334.
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=2287954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,

```

```

RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.M., Temple L., James K., Basham D., Bason N., Cherevach I.,
RA Achtmann M., Atkin R., Baker S., Cronin A., Davis P., Doggett J.,
RA Chillingworth T., Collins M., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moulé S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Saunders M., Saunders R., Squares S., Stevens K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Maskell D.J.;
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
RL EMBL; BX640447; CAE33826.1; -.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 217 AA; 23023 MW; 51158EA54B51856E CRC64;

Query Match 9.5%; Score 7; DB 16; Length 217;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IVAPALT 39
DB 170 IVAPALT 176
|||||
PRT; 217 AA.

RESULT 10
Q7WH915 PRELIMINARY;
ID Q7WH915 AC Q7WH915
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN BP1774.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=2287954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdeno-Tarraga A.M., Temple L., James K., Basham D., Bason N., Cherevach I.,
RA Achtmann M., Atkin R., Baker S., Cronin A., Davis P., Doggett J.,
RA Chillingworth T., Collins M., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moulé S., Norbertczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Saunders M., Saunders R., Squares S., Stevens K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Maskell D.J.;
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
RL EMBL; BX640428; CAE37075.1; -.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 217 AA; 22967 MW; 4108F2C4562DF46E CRC64;

Query Match 9.5%; Score 7; DB 16; Length 217;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IVAPALT 39
DB 170 IVAPALT 176
|||||
PRT; 228 AA.

RESULT 11
Q82UM7 PRELIMINARY;
ID Q82UM7 AC Q82UM7
DT 01-JUN-2003 (TrEMBLrel. 24, Created)

```

```

DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE GDSL lipolytic enzyme (EC 3.1.1.-).
CN NE1455.
OS Nitrosomonas europaea.
OC Bacteria; Proteobacteria; Betaproteobacteria; Nitrosomonadales;
OC Nitrosomonadaceae; Nitrosomonas.
OX NCBI_TaxID=915;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19718 / IFO 14298;
RX MEDLINE=22586410; PubMed=12700255;
RA Chain P., Lamerdin J.E., Larimer F.W., Regala W., Lao V., Land M.,
RA Hauser L., Hooper A.B., Klotz M.G., Norton J., Sayavedra-Soto L.A.,
RA Arclero D.M., Hommes N.G., Whittaker M.M., Arp D.J.;
RT "Complete genome sequence of the ammonia-oxidizing bacterium and
RT obligate chemolithoautotroph Nitrosomonas europaea.";
RL J. Bacteriol. 185:2759-2773 (2003).
DR EMBL; EX31861; CAD95366.1; -.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR001087; Lipase GDSL.
DR InterPro; IPR008265; Lipase GDSL_AS.
DR Pfam; PF00657; Lipase GDSL; 1.
DR PROSITE; PS01098; LIPASE GDSL_SER; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 228 AA; 24902 MW; 1290AF550781F5F7 CRC64;

Query Match          9.5%; Score 7; DB 16; Length 228;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
DB 199 PALTAVL 205

RESULT 12
Q92EW9 PRELIMINARY; PRT; 261 AA.
AC Q92EW9
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein lin0339.
CN LIN0339.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852 (2001).
DR EMBL; AL596164; CAC95572.1; -.
DR FIR; AD1475; AD1475.
DR ListList; LIN00339; -.
DR InterPro; IPR006879; YdJC.
DR Pfam; PF04794; YdJC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 261 AA; 29511 MW; 64D1AFD062DF46B4 CRC64;

Query Match          9.5%; Score 7; DB 16; Length 261;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 MPKLIID 28
DB 1 MPKLIID 7

RESULT 13
Q8YA48 PRELIMINARY; PRT; 261 AA.
AC Q8YA48
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Hypothetical protein lmo0314.
CN LMO0314.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852 (2001).
DR EMBL; AL591974; CAD00841.1; -.
DR FIR; AC1114; AC1114.
DR ListList; LMO0314; -.
DR InterPro; IPR006879; YdJC.
DR Pfam; PF04794; YdJC; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 261 AA; 29553 MW; 226B3DB837B87902 CRC64;

Query Match          9.5%; Score 7; DB 16; Length 261;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 MPKLIID 28
DB 1 MPKLIID 7

RESULT 14
Q8EC47 PRELIMINARY; PRT; 281 AA.
AC Q8EC47
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Transcriptional regulator, lysR family.
CN SO3297.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,

```

Job time : 16.0345 secs

RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Unayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A., C.M.,
RA Feldblum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.,
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015767; AAN56295.1; -
DR TIGR; SO3297; -
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR000847; HTH_LYSR.
DR Pfam; PF00126; HTH_1; 1.
DR PRINTS; PR00039; HTHLYSR.
DR PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW Complete proteome.
SQ SEQUENCE 281 AA; 32162 MW; 7781258C88463A7 CRC64;

Query Match 9.5%; Score 7; DB 16; Length 281;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 KLIIDTP 30
DB 206 KLIIDTP 212

RESULT 15

Q9YLM8 PRELIMINARY; PRT; 303 AA.
AC Q9YLM8;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein VP1 (Fragment).
GN VP1.
OS Human coxsackievirus A20.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=42782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IH-35;
RX MEDLINE=99138973; PubMed=99171773;
RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of
RT serotype with VP1 sequence and application to picornavirus
RT classification."
RL J. Virol. 73:1941-1948(1999).
DR EMBL; AF081309; AAD17713.1; -
DR HSSP; P03300; IPOV.
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00073; rhv; 1.
FT NON_TER 1
FT NON_TER 303
SQ SEQUENCE 303 AA; 33758 MW; 9DF80BA022B20A5C CRC64;

Query Match 9.5%; Score 7; DB 12; Length 303;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 APALTAV 41
DB 41 APALTAV 47

Search completed: March 14, 2004, 21:54:46

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:47:53 ; Search time 4.33793 Seconds
(without alignments)
888.256 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 74

Sequence: 1 ARAYXKMFMPFKAGKWC.....ARLXPCGKVGMDVWRRWS 74

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	7	9.5	442	1	GLNA_METH
2	7	9.5	469	1	GABP_BACSU
3	7	9.5	1067	1	LOM_SCHPO
4	7	9.5	2549	1	FRAP_HUMAN
5	7	9.5	2549	1	FRAP_MOUSE
6	7	9.5	2549	1	FRAP_RAT
7	6	8.1	106	1	ATPR_DROME
8	6	8.1	107	1	FER_BUCBP
9	6	8.1	109	1	NCYM_HUMAN
10	6	8.1	114	1	RSN_MOUSE
11	6	8.1	132	1	PRF_ANTGR
12	6	8.1	150	1	CYPH_ALLCE
13	6	8.1	155	1	BCT1_SHEEP
14	6	8.1	157	1	PR1_MEDSA
15	6	8.1	158	1	DRR4_PEA
16	6	8.1	160	1	SCS1_SHEEP
17	6	8.1	161	1	CUC1_TENMO
18	6	8.1	164	1	CYPH_BLAG
19	6	8.1	169	1	CYP2_ARATH
20	6	8.1	169	1	PP1B_HABIN
21	6	8.1	171	1	CYPH_BRANA
22	6	8.1	171	1	CYPH_LYCPS
23	6	8.1	172	1	CYPH_CATRO
24	6	8.1	172	1	CYPH_VAIZE
25	6	8.1	174	1	YGB0_YEAST
26	6	8.1	176	1	BCT5_BOVIN
27	6	8.1	183	1	AMPD_ECOLI
28	6	8.1	184	1	TR17_HUMAN
29	6	8.1	190	1	BCT7_SHEEP
30	6	8.1	198	1	REMO_SOLTU
31	6	8.1	221	1	YECA_ECOLI
32	6	8.1	225	1	LAPB_PASHA
33	6	8.1	226	1	Y374_HELPY

RESULT 1

ID	GLNA_METH	STANDARD	PRT	442 AA
AC	O27612			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	15-DEC-1998 (Rel. 37, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase) (GS).			
GN	GLNA OR MTH1570.			
OS	Methanobacterium thermoautotrophicum.			
OC	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;			
OC	Methanobacteriaceae; Methanothermobacter.			
OX	NCBI_TaxID=187420;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Delta H;			
RC	MEDLINE=98037514; PubMed=9371463;			
RA	Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,			
RA	Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,			
RA	Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,			
RA	Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,			
RA	Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,			
RA	McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,			
RA	Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;			
RT	"Complete genome sequence of Methanobacterium thermoautotrophicum			
RT	deltah: functional analysis and comparative genomics.";			
RL	J. Bacteriol. 179:7135-7155(1997).			
CC	-1- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +			
CC	L-glutamine.			
CC	-1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).			
CC	-1- SIMILARITY: Belongs to the glutamine synthetase family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AB000917; AAB86044.1; -			
DR	PIR; F69076; F69076.			
DR	HSP; P06201; ILGR.			
DR	InterPro; IPR008147; Gln_synt_beta.			
DR	InterPro; IPR008146; Gln_synt_C.			
DR	InterPro; IPR004809; GlnA.			
DR	Pfam; PF00120; gln-synt_1.			
DR	Pfam; PF03951; gln-synt_N; 1.			
DR	ProDom; PD001057; Gln_synt_C; 1.			
DR	TIGRFA; TIGR00653; GlnA; 1.			
DR	PROSITE; PS00180; GLNA_1; 1.			
DR	PROSITE; PS00181; GLNA_ATP; 1.			
KW	Ligase; Complete proteome.			
FT	372 372			
FT	AMP (UNDER CONDITIONS OF ABUNDANT			
FT	GLUTAMINE) (BY SIMILARITY).			
FT	SEQUENCE 442 AA; 50248 MW; B907928AAD8960A6 CRC64;			
FT	SEQUENCE 442 AA; 50248 MW; B907928AAD8960A6 CRC64;			

P42252 bacillus su
O9cno0 mycobacteri
O51575 borrelia bu
P48296 thermus aqu
Q82330 pyrobaculum
Q8d4q0 vibrio vuln
Q9y3p4 homo sapien
P35487 mus musculu
P57061 neisseria m
P57062 neisseria m
O04556 arabidopsis
O04558 arabidopsis

```

Query Match          9.5%; Score 7; DB 1; Length 442;
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 APALTAV 41
Db 281 APALTAV 287

RESULT 2
GABP_BACSU STANDARD; PRT; 469 AA.
ID P46349; P94473;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE GABA permease (4-amino butyrate transport carrier) (Gamma-
DE aminobutyrate permease).
GN GABP OR BSU06310.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97109536; PubMed=8951816;
RA Person A.E., Gray L.V., Fisher S.H.;
RT "Expression of the Bacillus subtilis gabp gene is regulated
RT independently in response to nitrogen and amino acid availability.";
RL Mol. Microbiol. 22:693-701(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=97124186; PubMed=8969499;
RA Borries R., Porwollik S., Schroeter R.;
RT "The 52 degrees-55 degrees segment of the Bacillus subtilis
RT chromosome: a region devoted to purine uptake and metabolism, and
RT containing the genes cotA, gabp and guaA and the pur gene cluster
RT within a 34560 bp nucleotide sequence.";
RL Microbiology 142:3027-3031(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Besterio M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerion I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Fertari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
RA Gim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haeche J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerz-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidis E., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Mauda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,
RA Neone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Prescann E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sakiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,

```

The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*,"
 Nature 390:249-256(1997).
 CC -!- FUNCTION: TRANSPORTER FOR GABA.
 CC -!- PATHWAY: 4-aminobutyrate (GABA) degradation.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the amino acid permease family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U31756; AAC44641.1; -.
 CC EMBL; U51115; AAB62306.1; -.
 CC EMBL; Z99107; CAB12450.1; -.
 CC PIR; C69628; C69628.
 CC Subtilist; BGL1328; gabp.
 CC InterPro; IPR002293; AA/rel_permease1.
 CC InterPro; IPR003422; AA/rel_permease2.
 CC InterPro; IPR004840; AAC_permease.
 CC InterPro; IPR004841; Permease region.
 CC Pfam; PF00324; aa_permeases; 1.
 CC PROSITE; PS00218; AMINO ACID PERMEASE_1; 1.
 CC Transport; Amino-acid transport; Transmembrane; Complete proteome.
 KW TRANSMEM 18 38
 FT TRANSMEM 40 60 POTENTIAL.
 FT TRANSMEM 95 115 POTENTIAL.
 FT TRANSMEM 117 137 POTENTIAL.
 FT TRANSMEM 158 178 POTENTIAL.
 FT TRANSMEM 201 221 POTENTIAL.
 FT TRANSMEM 243 263 POTENTIAL.
 FT TRANSMEM 270 290 POTENTIAL.
 FT TRANSMEM 329 349 POTENTIAL.
 FT TRANSMEM 356 376 POTENTIAL.
 FT TRANSMEM 402 422 POTENTIAL.
 FT TRANSMEM 427 447 POTENTIAL.
 FT CONFLICT 36 36 P -> S (IN REF. 1).
 SQ SEQUENCE 469 AA; 51094 MW; 41900776678F0B62 CRC64;

Query Match 9.5%; Score 7; DB 1; Length 469;
 Best Local Similarity 100.0%; Pred. No. 8.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 LTAVLSC 44
 Db 291 LTAVLSC 297
 |||||
 |||||

RESULT 3
 LONM_SCHPO STANDARD; PRT; 1067 AA.
 ID LONM_SCHPO
 AC Q09769;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Putative lon protease homolog, mitochondrial precursor (EC 3.4.21.-).
 GN SPAC22F3.06C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson B., Quail M.A., Rabinowitz E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weljens J., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Frizc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillard C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsberg S.L.,
 RA Cerrutti L., Lowe I., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RT "The genome sequence of Schizosaccharomyces pombe."
 RL Nature 415:871-880(2002).
 CC -!- FUNCTION: Required for intramitochondrial proteolysis. Catalyzes
 CC the initial steps of protein degradation (Potential).
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix (Potential).
 CC -!- SIMILARITY: Belongs to peptidase family S16.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; Z54285; CAA91071.1; --
 DR PIR; S62421; S62421.
 DR MEROPS; S16.002;
 DR GenDB_Spombe; SPAC22F3.06c; --
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR003959; AAA ATPase centr.
 DR InterPro; IPR008269; Pept_S16_C.
 DR InterPro; IPR004815; Pept_S16_Lon.
 DR InterPro; IPR003111; Pept_S16_N.
 DR InterPro; IPR008268; Peptid_S16_AS.
 DR InterPro; IPR001994; Peptidase_S16.
 DR Pfam; PF00004; AAA; 1.
 DR Pfam; PF02190; LON; 1.
 DR Pfam; PF05362; Lon C; 1.
 DR PRINTS; PR00830; ENDOLATASE.
 DR SMART; SM00382; AAA; 1.
 DR SMART; SM00464; LON; 1.
 DR TIGRFAMs; TIGR00763; lon; 1.
 DR PROSITE; PS01046; LON_SER; 1.
 DR KW Hypothetical protein; Hydroxylase; Serine protease; ATP-binding;
 KW Mitochondrion; Transist peptide.
 FT TRANSIT 1 ? MITOCHONDRION (POTENTIAL).
 FT CHAIN ? 1067 PUTATIVE LON PROTEASE HOMOLOG.
 FT NFBIND 578 585 ATP (POTENTIAL).
 FT ACT_SITE 946 946 BY SIMILARITY.
 SQ SEQUENCE 1067 AA; 118641 MW; 215FCFBE9CDC4AAD CRC64;

Query Match 9.5%; Score 7; DB 1; Length 1067;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
 Db 244 PALTAVL 250

RESULT 4

FRAP HUMAN
 ID FRAP HUMAN STANDARD; PRT; 2549 AA.
 AC P42345; O9Y4I3;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE FRAP1-rapamycin associated protein (FRAP) (Rapamycin target protein).
 GN FRAP1 OR FRAP OR FRAP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=94277209; PubMed=8008069;
 RA Brown E.J., Albers M.W., Shin T.B., Ichikawa K., Keith C.T.,
 RA Lane W.S., Schreiber S.L.;
 RT "A mammalian protein targeted by G1-arresting rapamycin-receptor
 RT complex."
 RL Nature 369:756-758(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98317532; PubMed=9653645;
 RA Ouyang P., Lubyova B., Gardellin P., Kurzbaue R., Weith A.;
 RT "Molecular cloning and expression analysis of five novel genes in
 RT chromosome 1p36."
 RL Genomics 50:187-198(1998).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 2018-2112.
 RX MEDLINE=96279639; PubMed=8662507;
 RA Choi J., Chen J., Schreiber S.L.; Clardy J.;
 RT "Structure of the FKBP12-rapamycin complex interacting with the
 RT binding domain of human FRAP."
 RL Science 273:239-242(1996).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 2018-2112.
 RX MEDLINE=99190960; PubMed=10089303;
 RA Liang J., Choi J., Clardy J.;
 RT "Refined structure of the FKBP12-rapamycin-FRB ternary complex at 2.2
 RT A resolution."
 RL Acta Crystallogr. D 55:736-744(1999).
 CC -!- FUNCTION: ACTS AS THE TARGET FOR THE CELL-CYCLE ARREST AND
 CC IMMUNOSUPPRESSIVE EFFECTS OF THE FKBP12-RAPAMYCIN COMPLEX.
 CC -!- SIMILARITY: Belongs to the P13/P14-kinase family.
 CC -!- SIMILARITY: Contains 8 HEAT repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L34075; AAA58486.1; --
 DR EMBL; U88966; AAC39933.1; --
 DR PIR; S45340; S45340.
 DR PDB; 1FAP; 23-JUL-97.
 DR PDB; 1NSG; 18-MAR-98.
 DR PDB; 1AUE; 18-NOV-98.
 DR PDB; 2FAP; 09-AUG-99.
 DR PDB; 3FAP; 13-SEP-00.
 DR PDB; 4FAP; 13-SEP-00.
 DR Genew; HGNC:3942; FRAP1.
 DR MM; 601231; --
 DR GO; GO:0000074; P:regulation of cell cycle; TAS.
 DR InterPro; IPR008939; ARM.
 DR InterPro; IPR003151; FAT.
 DR InterPro; IPR003152; FATC.
 DR InterPro; IPR000357; HEAT.
 DR InterPro; IPR000403; P13_P14_kinase.
 DR InterPro; IPR008940; Prenyl_Trans.


```

RX MEDLINE=94306515; PubMed=7518356;
RA Sabatini D.M., Erdjument-Bromage H., Lui M., Tempet P.,
RA Snyder S.H.;
RT "RAP11: a mammalian protein that binds to FKBP12 in a rapamycin-
RT dependent fashion and is homologous to yeast TORB.";
RL Cell 78:35-43(1994).
CC -!- FUNCTION: ACTS AS THE TARGET FOR THE CELL-CYCLE ARREST AND
CC IMMUNOSUPPRESSIVE EFFECTS OF THE FKBP12-RAPAMYCIN COMPLEX.
CC -!- SIMILARITY: Belongs to the P13/P14-kinase family.
CC -!- SIMILARITY: Contains 8 HEAT repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: L37085; AAA65929.1; -.
CC DR EMBL: U11681; AAA20091.1; -.
CC DR PIR: A54837; A54837.
CC DR HSSP: P42345; 1FAP.
CC DR InterPro: IPR008938; ARM.
CC DR InterPro: IPR003151; FAT.
CC DR InterPro: IPR003152; FATC.
CC DR InterPro: IPR000357; HEAT.
CC DR InterPro: IPR004033; P13_P14_kinase.
CC DR InterPro: IPR008940; Ptenyl_trans.
CC DR Pfam: PF02259; FAT_1.
CC DR Pfam: PF02260; FATC_1.
CC DR Pfam: PF00454; P13_P14_kinase; 1.
CC DR SMART: SM00146; PI3K; 1.
CC DR PROSITE: PS00915; P13_4_KINASE_1; 1.
CC DR PROSITE: PS00916; P13_4_KINASE_2; 1.
CC DR PROSITE: PS0290; P13_4_KINASE_3; 1.
CC DR PROSITE: PS00077; HEAT_REPEAT_FALSE_NEG.
CC DR Transferrase; Kinase; Repeat.
KW REPEAT 16 53
FT REPEAT 650 688 HEAT 1.
FT REPEAT 859 897 HEAT 2.
FT REPEAT 988 1025 HEAT 3.
FT REPEAT 1069 1106 HEAT 4.
FT REPEAT 1109 1148 HEAT 5.
FT REPEAT 1150 1186 HEAT 6.
FT REPEAT 1382 1392 HEAT 7.
FT DOMAIN 1393 1370 FAT.
FT REPEAT 2182 2549 PI3K/P14K.
FT DOMAIN 2182 2549
SQ SEQUENCE 2549 AA; 288791 MW; BE841EA7B9086F99 CRC64;

Query Match 9.5%; Score 7; DB 1; Length 2549;
Best Local Similarity 100.0%; Fred.No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 PALTAVL 42
DB 515 PALTAVL 521
|||||

RESULT 7
ID ATPR DROME STANDARD; PRT; 106 AA.
AC Q24407; Q9VCN0;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ATP synthase coupling factor 5, mitochondrial precursor (EC 3.6.3.14)
(GF6).
GN APTSYN-CF6 OR CG4412.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.

```

```

OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=99168769; PubMed=10071211;
RA Caggese C., Ragone G., Ferrini B., Moschetti R., De Pinto V.,
RA Caizzi R., Barsanti P.;
RT "Identification of a collection of D. melanogaster cDNAs homologous to
RT sequences in the Human Gene Index database.";
RL Mol. Gen. Genet. 261:64-70(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Adamatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.B.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Ffannkoch C., Baldwin D.,
RA Beeson K.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Berkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Galbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirekas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: This is one of the chains of the nonenzymatic component
CC (CF0) subunit of the mitochondrial ATPase complex. F6 seems to
CC be part of the stalk that links CF(0) to CF(1).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -!- SUBUNIT: F-type ATPases have 2 components, CF(1) - the catalytic
CC core - and CF(0) - the membrane proton channel. CF(0) seems to
CC have nine subunits: a, b, c, d, e, f, g, F6 and 8 (or A6L).
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL: X99665; CAA67979.1; -.
CC DR EMBL: A5003743; AAF56127.1; -.
CC DR FlyBase; FBgn0016119; APTsyn-Cf6.

```

```

DR InterPro; IPR008387; ATP synth_F6.
KW Pfam; PF05111; ATP-synt_F6; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transit peptide.
FT TRANSIT 1 2 MITOCHONDRION (POTENTIAL)
FT CHAIN 7 106 ATP SYNTHASE COUPLING FACTOR 6.
SQ SEQUENCE 106 AA; 11936 MW; 08E1B074EB34E94B CRC64;

Query Match 8.1%; Score 6; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IVAPAL 38
DB 23 IVAPAL 28

RESULT 8
PER_BUCBP STANDARD; PRT; 107 AA.
AC Q89A15;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Ferredoxin, 2Fe-2S.
CN PDX OR BFP548
OS Buchnera aphidicola (subsp. Baizongia pistaciae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Buchnera.
OX NCBI_TaxID=135842;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426901; PubMed=12522365;
RA Van Ham R.C.H.U.; Kamerbeek J.; Palacios C.; Rausell C.; Abascal F.;
RA Bastolla U.; Fernandez J.M.; Jimenez L.; Postigo M.; Silva F.J.;
RA Tamames J.; Viguera E.; Latorre A.; Valencia A.; Moran F.; Moya A.;
RT "Reductive genome evolution in Buchnera aphidicola.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:581-586(2003).
CC -!- FUNCTION: FERREDOXIN ARE IRON-SULFUR PROTEINS THAT TRANSFER
CC -!- ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.
CC -!- COFACTOR: Binds 1 2Fe-2S cluster (by similarity).
CC -!- SIMILARITY: Belongs to the adrenodoxin / putidaredoxin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; A8014017; AAC27246.1; -.
DR InterPro; IPR001035; Adrenodoxin.
DR Pfam; PF001041; Ferredoxin.
DR PRINTS; PR00111; fer2; 1.
DR PROSITE; PS00814; ADX; 1.
KW Electron transport; Metal-binding; Iron-sulfur; Iron; 2Fe-2S;
KW Complete proteome.
FT METAL 42 42 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 48 48 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 51 51 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
FT METAL 87 87 IRON-SULFUR (2FE-2S) (BY SIMILARITY).
SQ SEQUENCE 107 AA; 12118 MW; 47A2E9FC92B1D649 CRC64;

Query Match 8.1%; Score 6; DB 1; Length 107;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 MPKLII 27
DB 1 MPKLII 6

DR InterPro; IPR008387; ATP synth_F6.
KW Pfam; PF05111; ATP-synt_F6; 1.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transit peptide.
FT TRANSIT 1 2 MITOCHONDRION (POTENTIAL)
FT CHAIN 7 106 ATP SYNTHASE COUPLING FACTOR 6.
SQ SEQUENCE 106 AA; 11936 MW; 08E1B074EB34E94B CRC64;

Query Match 8.1%; Score 6; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 IVAPAL 38
DB 23 IVAPAL 28

RESULT 9
NCYM HUMAN STANDARD; PRT; 109 AA.
AC P40205;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE N-cym protein.
GN NCYM OR CYMN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93041371; PubMed=1419902;
RA Armstrong B.C.; Krystal G.W.;
RT "Isolation and characterization of complementary DNA for N-cym, a
RL gene encoded by the DNA strand opposite to N-myc.";
RL Cell Growth Differ. 3:385-390(1992).
CC -!- FUNCTION: May have a functional role during normal fetal
CC development.
CC -!- TISSUE SPECIFICITY: Expressed in the fetal brain, lung, liver and
CC kidney at varying low levels.
CC -!- DEVELOPMENTAL STAGE: Expressed during fetal development, as well
CC as in tumor cell lines containing amplified N-myc loci, where it
CC is expressed at very high levels.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; S49953; AAB24206.1; -.
DR PIR; A49009; A49009.
DR MIM; 605374; -.
DR GO; GO:0007275; P:development; TAS.
FT DOMAIN 68 109 ARG/LYS-RICH (BASIC).
SQ SEQUENCE 109 AA; 11746 MW; E399C428CF07F6E7 CRC64;

Query Match 8.1%; Score 6; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 APALTA 40
DB 38 APALTA 43

RESULT 10
RSN_MOUSE STANDARD; PRT; 114 AA.
AC Q99F87;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Resistin precursor (Cysteine-rich secreted protein FIZ23) (Adipose
DE tissue-specific secretory factor) (ADSF) (Adipose-specific cysteine-
DE rich secreted protein A12-alpha).
GN RETN OR FIZ23.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.; PARTIAL SEQUENCE, AND CHARACTERIZATION.
RX MEDLINE=21069045; PubMed=11201732;
RA Steppan C.M.; Bailey S.T.; Bhat S.; Brown E.J.; Banerjee R.R.;
RA Wright C.M.; Patel H.R.; Ahima R.S.; Lazar M.A.;
RT "The hormone resistin links obesity to diabetes.";

```

RL Nature 409:307-312 (2001).

RN [2]

RP SEQUENCE FROM N.A.

RA Rajala M.W., Scherer P.E.;

RT "Identification of a novel adipose-specific cysteine-rich secreted

RT protein."

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Mammary gland;

RA Strausberg R.L., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Klausner R.D., Feingold E.A., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Altschul S.F., Jordan H., Moore T., Max S.I., Wang J., Haeh P.,

RA Hopkins R.F., Maruina K., Farmer A.A., Rubin G.M., Hong L.,

RA Diatchenko L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loguillano N.A., Peters G.J., Abramson R.D., Mullany S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Faney J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences."

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [4]

RP SUBUNIT.

RX MEDLINE=21336653; PubMed=11358969;

RA Banerjee R.R., Lazar M.A.;

RT "Dimerization of resistin and resistin-like molecules is determined by

RT a single cysteine."

RL J. Biol. Chem. 276:25970-25973 (2001).

CC -!- FUNCTION: Hormone that seems to suppress insulin ability to

CC stimulate glucose uptake into adipose cells. Potentially

CC links obesity to diabetes.

CC -!- SUBUNIT: Homodimer; disulfide-linked.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: Expressed in white but not brown adipose

CC tissue in a variety of organs.

CC -!- PTM: 5 disulfide bonds are present (Probable).

CC -!- SIMILARITY: Belongs to the resistin/FIZZ family.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC EMBL: AF223080; AK59823.1; -

CC EMBL: AF220870; AK3102.1; -

CC EMBL: BC051196; AAH51196.1; -

CC MGD: MG1:1888506; Refn.

CC DR GO: GO:0005576; C:extracellular; IDA.

CC KW Hormone; Signal; Diabetes mellitus; Obesity.

FT SIGNAL 1 20

FT CHAIN 21 114 RESISTIN

FT DISULFID 26 26 INTERCHAIN (PROBABLE).

CC SEQUENCE 114 AA; 12491 MW; D4930E51D3F22C8 CRC64;

Query Match 8.1%; Score 6; DB 1; Length 114;

Best Local Similarity 100.0%; Pred. No. 32;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 TAVLSC 44

|||||

Db 71 TAVLSC 76

RESULT 11

PRT_ANTGR STANDARD; PRT; 132 AA.

ID PRT_ANTGR

AC P17502;

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-FEB-1991 (Rel. 17, Last annotation update)

DE Protamine.

OS Anthomus grandis (Boll weevil).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;

OC Phytophaga; Curculionidae; Curculioninae; Anthonomini; Anthomus.

OX NCBI_TaxID=7044;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RX MEDLINE=90301495; PubMed=2362819;

RA Trewitt P.M., Heilmann L.J., Kumaran A.K.;

RT "Boll weevil testis-specific cDNA."

RL Nucleic Acids Res. 18:3646-3646 (1990).

CC -!- FUNCTION: Protamines substitute for histones in the chromatin of

CC sperm during the haploid phase of spermatogenesis. They compact

CC sperm DNA into a highly condensed, stable and inactive complex.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- TISSUE SPECIFICITY: Testis.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

CC EMBL: X52058; CAA36282.1; -

CC EMBL: S10305; S10305.

CC Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;

CC Testis; DNA condensation; Nuclear protein.

CC SEQUENCE 132 AA; 16280 MW; 8AC65334FCE300D CRC64;

Query Match 8.1%; Score 6; DB 1; Length 132;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 KCWCKM 22

|||||

Db 77 KCWCKM 82

RESULT 12

CYPH_ALICE STANDARD; PRT; 150 AA.

ID CYPH_ALICE

AC P34887;

DT 01-FEB-1994 (Rel. 28, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Peptidyl-prolyl cis-trans isomerase (EC 5.2.1.8) (PPIase) (Rotamase)

DE (Cyclophilin) (Cyclosporin A-binding protein).

GN CYP.

OS Allium cepa (Onion).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Alliaceae;

OC Allium.

OX NCBI_TaxID=4679;

RN [1]

RP SEQUENCE FROM N.A.

RA Clark S.A.;

RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: PPIases accelerate the folding of proteins. It catalyzes

CC the cis-trans isomerization of proline imidic peptide bonds in

CC oligopeptides.
 CC -1- CATALYTIC ACTIVITY: Peptidylproline (omega=180) = peptidylproline
 CC (Omega=0).
 CC -1- ENZYME REGULATION: Binds cyclosporin A (CSA). CSA mediates some of
 CC its effects via an inhibitory action on PPIase.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the cyclophilin-type PPIase family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L13365; AAA32642.1; --
 CC HSP; P05092; 2CPL.
 CC InterPro; IPR002130; CSA_PPIase.
 CC Pfam; PF00160; pro_isomerase; 1.
 CC PRINTS; PR00153; CSAPPISMPASE.
 CC PROSITE; PS00170; CSA_PPIASE_1; 1.
 CC PROSITE; PS00072; CSA_PPIASE_2; 1.
 CC Cyclosporin; Isomerase; Rotamase; Multigene family.
 CC SEQUENCE 150 AA; 16033 MW; 9223D16840F7E241 CRC64;
 CC -----
 CC Query Match 8.1%; Score 6; DB 1; Length 150;
 CC Best Local Similarity 100.0%; Pred.No. 40;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 63 VEGMDV 68
 CC DB 118 VEGMDV 123
 CC | | | | |
 CC -----
 CC RESULT 13
 CC BCTL SHEEP
 CC ID BCTL SHEEP STANDARD; PRT; 155 AA.
 CC AC P54230;
 CC DT 01-OCT-1996 (Rel. 34, Created)
 CC DT 01-OCT-1996 (Rel. 34, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Cyclic dodecapeptide precursor (Bactenecin 1).
 CC GN DODEA AND DODEB
 CC OS Ovis aries (Sheep).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 CC OC Bovidae; Caprinae; Ovis.
 CC OX NCBI_TaxID=9940;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Bone marrow;
 CC RX MEDLINE=96105386; PubMed=7498547;
 CC RA Bagella L., Scocchi M., Zanetti M.;
 CC FT "cDNA sequences of three sheep myeloid cathelicidins.";
 CC RL FEBS Lett. 376:225-228(1995).
 CC RN [2]
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Liver;
 CC RA Hutter K.M., Mahoney M.M.;
 CC RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Potent microbicidal activity, active against
 CC Staphylococcus aureus and Escherichia coli (By similarity).
 CC -1- SIMILARITY: Belongs to the cathelicidin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L46853; AAA85469.1; --
 DR EMBL; U60595; AAB49710.1; --
 DR EMBL; U60596; AAB49711.1; --
 DR PIR; S68229; S69229.
 DR InterPro; IPR001894; Cathelicidin.
 DR Pfam; PF00666; Cathelicidins; 1.
 DR PRODOM; P0001836; Cathelicidin; 1.
 DR PROSITE; PS00946; Cathelicidins_1; 1.
 DR PROSITE; PS00947; Cathelicidins_2; 1.
 KW Antibiotic; Signal; Pyrrolidone carboxylic acid.
 KW Antibiocytic; Signal; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 29 POTENTIAL.
 FT PROPEP 30 143 BY SIMILARITY.
 FT PEPTIDE 144 155 CYCLIC DODECAPEPTIDE.
 FT MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID
 FT (BY SIMILARITY).
 FT DISULFID 85 96 BY SIMILARITY.
 FT DISULFID 107 124 BY SIMILARITY.
 FT DISULFID 146 154 BY SIMILARITY.
 SQ SEQUENCE 155 AA; 17648 MW; 1690638C791B1736 CRC64;
 CC -----
 CC Query Match 8.1%; Score 6; DB 1; Length 155;
 CC Best Local Similarity 100.0%; Pred.No. 41;
 CC Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 47 RCSLWL 52
 CC DB 12 RCSLWL 17
 CC | | | | |
 CC -----
 CC RESULT 14
 CC PRI MEDSA
 CC ID PRI MEDSA STANDARD; PRT; 157 AA.
 CC AC Q43560;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Class-10 pathogenesis-related protein 1 (MSPR10-1).
 CC GN MSPR10-1.
 CC OS Medicago sativa (Alfalfa).
 CC OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 CC OC eurosid1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
 CC OX NCBI_TaxID=3879;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC STRAIN=cv. Nagyszénási; TISSUE=Leaf;
 CC MEDLINE=97023949; PubMed=8870270;
 CC Breda C., Sallaud C., El-Turk J., Buffard D., De Kozak I., Banault R.,
 CC Kondorosi A.;
 CC RT "Defense reaction in Medicago sativa: a gene encoding a class 10 PR
 CC protein is expressed in vascular bundles.";
 CC RL Mol. Plant Microbe Interact. 9:713-719(1996).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: High levels in roots and not detectable in
 CC hypocotyle, cotyledons, stems, leaves and flower buds of untreated
 CC plants. After induction, high levels are present in the vascular
 CC bundles of leaves.
 CC -1- DEVELOPMENTAL STAGE: It appears in leaves 3 hours after induction,
 CC maximum levels are reached by 24 hours and remain at a high
 CC level over a period of at least 72 hours.
 CC -1- INDUCTION: Upon contact with the plant pathogen Pseudomonas
 CC syringae pv. pisi and salicylic acid.
 CC -1- SIMILARITY: Belongs to the BctVI family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X98867; CAA67375.1; --
 CC -----

DR PIR: T09659; T09659.
DR HSP: P15494; 1BV1.
DR InterPro: IPR000916; Bet v I.
DR Pfam: PF00407; Bet v I; I.
DR PRINTS: PR00634; BETALLERGEN.
DR ProDom: PD000531; Bet v I; 1.
DR PROSITE: PS00451; PATHOGENESIS BETVI; 1.
KW Plant defense; Pathogenesis-related protein; Multigene family.
SQ SEQUENCE 157 AA; 16645 MW; 635E986EF0BC4BFC CRC64;

Query Match 8.1%; Score 6; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 SIVAPA 37
Db 12 SIVAPA 17
|||||

RESULT 15
ID_DRR4_PEA STANDARD; PRT; 158 AA.
AC P27037;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Disease resistance response protein DRRG49-C.
OS Pisum sativum (Garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids 1; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Alaska;
EX MEDLINE=92063198; PubMed=2132027;
RA Chiang C.C.; Hadwiger L.A.;
RT "Cloning and characterization of a disease resistance response gene
in pea inducible by Fusarium solani.";
RL Mol. Plant Microbe Interact. 3:78-85(1990).
CC -!- INDUCTION: Upon contact with the plant pathogens fungus Fusarium
CC solani.
CC -!- SIMILARITY: Belongs to the BetVI family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J03680; AAA33663.1; -
DR HSP: P15494; 1BV1.
DR InterPro: IPR000916; Bet v I.
DR Pfam: PF00407; Bet v I; I.
DR PRINTS: PR00634; BETALLERGEN.
DR ProDom: PD000531; Bet v I; 1.
DR PROSITE: PS00451; PATHOGENESIS BETVI; 1.
KW Plant defense; Pathogenesis-related protein; Multigene family.
SQ SEQUENCE 158 AA; 16791 MW; F7F5E87B52D1A4DC CRC64;

Query Match 8.1%; Score 6; DB 1; Length 158;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 SIVAPA 37
Db 12 SIVAPA 17
|||||

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:49:49 ; Search time 7.4 Seconds
(without alignments)
961.915 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 74
Sequence: 1 ARAYXKMFNFMEKAGKCWC.....APLXPGKVGMDVWRRRWS 74

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8	10.8	2338	2	I73957
2	7	9.5	98	2	G96029
3	7	9.5	261	2	AC1114
4	7	9.5	261	2	AD1475
5	7	9.5	312	2	T35887
6	7	9.5	434	2	H70989
7	7	9.5	442	2	F69076
8	7	9.5	444	2	T22412
9	7	9.5	469	2	C69628
10	7	9.5	537	2	AG3572
11	7	9.5	1067	1	S62421
12	7	9.5	2549	2	S45340
13	7	9.5	2549	2	AS4837
14	6	8.1	60	2	C81686
15	6	8.1	70	2	D75413
16	6	8.1	79	2	AB2175
17	6	8.1	84	2	E69293
18	6	8.1	101	2	S07706
19	6	8.1	107	2	G97296
20	6	8.1	109	2	A49009
21	6	8.1	116	2	F81796
22	6	8.1	118	2	T05812
23	6	8.1	119	2	A96508
24	6	8.1	132	2	S10305
25	6	8.1	139	2	AE2921
26	6	8.1	141	2	S31702
27	6	8.1	146	2	T47310
28	6	8.1	152	2	S68411
29	6	8.1	152	2	B97184

antimicrobial pept
pathogenesis-relat
conserved hypothet
stress response ge
pathogenesis-relat
peptidylprolyl iso
peptidylprolyl iso
peptidylprolyl ci
peptidylprolyl ci
peptidylprolyl ci
peptidylprolyl ci
hypothetical prote
peptidylprolyl iso
peptidylprolyl iso
peptidylprolyl iso
peptidylprolyl iso
peptidylprolyl iso
peptidylprolyl iso

ALIGNMENTS

RESULT 1

I73957
kinase-related protein c-ros-1 precursor - rat
N:Contains: protein-tyrosine kinase (EC 2.7.1.112) ros-1
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 02-Jun-2000
C:Accession: I73957; I56752; I73956
R:Matsushime, H.; Shibuya, M.
J. Virol. 64, 2117-2125, 1990
A:Title: Tissue-specific expression of rat c-ros-1 gene and partial structural similarit
A:Reference number: I56752; MUID:90219211; PMID:2139140
A:Accession: I73957
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-2338 <RES>
A:Cross-references: GB:M35106; NID:G203599; PIDN:AAA40968.1; PID:G203600
A:Accession: I56752
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-430,452-2338 <RE2>
A:Cross-references: GB:M35104; NID:G203595; PIDN:AAA40966.1; PID:G203596
A:Accession: I73956
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-430,452-1872 'AC', 1875 <RE3>
A:Cross-references: GB:M35105; NID:G203597; PIDN:AAA40967.1; PID:G203598
C:Superfamily: Kinase-related protein ros; LDL receptor YWTD-containing repeat homology;
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; Kinase-related
ic protein kinase
F:753-793/Domain: LDL receptor YWTD-containing repeat homology <YW3>
F:1935-2314/Domain: protein kinase homology <KIN>
F:1943-1951/Region: protein kinase ATP-binding motif

Query Match 10.8%; Score 8; DB 2; Length 2338;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 29 TPFSIVAP 36
DB 413 TPFSIVAP 420

RESULT 2
G96029
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasamid pSymB
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: G96029
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasamid from the N2-fixing endo
A:Reference number: A95842; MUID:21396508; PMID:11481431

A:Accession: G96029
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-98 <KOR>
 A:Cross-references: GB:AL591985; PIDN:CAC49903.1; PID:gl5141391; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, H.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMb20591
 A:Genome: plasmid

Query Match 9.5%; Score 7; DB 2; Length 98;
 Best Local Similarity 100.0%; Pred. No. 4.6;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 38 LTAVLSC 44
 |||||
 Db 18 LTAVLSC 24

RESULT 3
 AC1114
 hypothetical protein lmo0314 [imported] - *Listeria monocytogenes* (strain EGD-e)
 C:Species: *Listeria monocytogenes*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AC1114
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Meok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
 A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AC1114
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-261 <GLA>
 A:Cross-references: GB:NC_003210; PIDN:CAD00841.1; PID:gl6409678; GSPDB:GN00177
 A:Experimental source: strain EGD-e
 C:Genetics:
 A:Gene: lmo0314

Query Match 9.5%; Score 7; DB 2; Length 261;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 MPKLIID 28
 |||||
 Db 1 MPKLIID 7

RESULT 4
 AD1475
 hypothetical protein homolog lin0339 [imported] - *Listeria innocua* (strain Clip11262)
 C:Species: *Listeria innocua*
 C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
 R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
 Science 294, 849-852, 2001
 A:Authors: Krett, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Meok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669
 A:Accession: AD1475
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-261 <GLA>
 A:Cross-references: GB:AL592022; PIDN:CAC95572.1; PID:gl6412768; GSPDB:GN00178
 A:Experimental source: strain Clip11262
 C:Genetics:
 A:Gene: lin0339

Query Match 9.5%; Score 7; DB 2; Length 261;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 MPKLIID 28
 |||||
 Db 1 MPKLIID 7

RESULT 5
 T35887
 hypothetical protein SC9B10.18 - *Streptomyces coelicolor*
 C:Species: *Streptomyces coelicolor*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
 C:Accession: T35887
 R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, November 1997
 A:Reference number: Z21592
 A:Accession: T35887
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-312 <OLI>
 A:Cross-references: EMBL:AL009204; PIDN:CAA15808.1; GSPDB:GN00070; SCOEDB:SC9B10.18
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC9B10.18

Query Match 9.5%; Score 7; DB 2; Length 312;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 35 APALTAV 41
 |||||
 Db 214 APALTAV 220

RESULT 6
 H70989
 probable GABA permease - *Mycobacterium tuberculosis* (strain H37RV)
 N:Alternate names: 4-amino butyrate transport carrier
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 21-Nov-1998
 C:Accession: H70989
 R:Col, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: H70989
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-434 <COL>
 A:Cross-references: GB:Z97831; GB:AL123456; NID:g3261825
 A:Experimental source: strain H37RV
 A>Note: this ORF is not annotated in GenBank entry MT20G10, release 106
 C:Genetics:
 A:Gene: Rv0522; gabP
 C:Superfamily: arginine permease

Query Match 9.5%; Score 7; DB 2; Length 434;
 Best Local Similarity 100.0%; Pred. No. 15;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 LTAVLSC 44
|||||

DB 271 LTAVLSC 277
|||||

RESULT 7

F99076

Glutamine synthetase - Methanobacterium thermoautotrophicum (strain Delta H)

C:Species: Methanobacterium thermoautotrophicum

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 16-Jul-1999

C:Accession: F69076

R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N. K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func01

A:Reference number: A69000; PMID:98037514; PMID:9371463

A:Accession: F69076

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-442 <MTH>

A:Cross-references: GB:AE000017; GB:AE000666; NID:G2622689; PIDN:AA386044.1; PID:G262269

A:Experimental source: strain Delta H

C:Genetics:

A:Gene: MTH1570

C:Superfamily: glutamate-ammonia ligase

Query Match 9.5%; Score 7; DB 2; Length 442;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 35 APALTAV 41
|||||

DB 281 APALTAV 287
|||||

RESULT 8

T22412

hypothetical protein F49C12.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22412

R:Gardner, A.

submitted to the EMBL Data Library, December 1995

A:Reference number: Z19562

A:Accession: T22412

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-444 <WIL>

A:Cross-references: EMBL:266227; PIDN:CAA92511.1; GSPDB:GN00022; CESP:F49C12.7

A:Experimental source: clone F49C12

C:Genetics:

A:Gene: CESP:F49C12.7

A:Map position: 4

A:Introns: 24/2; 78/3; 117/1; 144/3; 324/3; 380/3; 405/3

Query Match 9.5%; Score 7; DB 2; Length 444;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
|||||

DB 398 PALTAVL 404
|||||

RESULT 9

C69628

gamma-aminobutyrate permease gabb - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000

C:Accession: C69628

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte: C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chk A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Puma, S.; Galizzi, A.; Galle: iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauesl Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle Rieger, M.; Rivolta, C.; Rocca, E.; Roche, M.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seroi akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wibat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A:Authors: Yoshikawa, H.F.; Zumschein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; PMID:98044033; PMID:9384377

A:Accession: C69628

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-469 <KUN>

A:Cross-references: GB:Z99107; GB:AL009126; NID:G2632866; PIDN:CAB12450.1; PID:G2632944

A:Experimental source: strain 168

C:Genetics:

A:Gene: gabb

C:Superfamily: arginine permease

Query Match 9.5%; Score 7; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 38 LTAVLSC 44
|||||

DB 291 LTAVLSC 297
|||||

RESULT 10

AG3572

oligo peptide-binding protein appA precursor [imported] - Brucella melitensis (strain 16)

C:Species: Brucella melitensis

C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C:Accession: AG3572

R:DelVecchio, V.G.; Kapral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, M.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002

A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis

A:Reference number: AD3252; PMID:11756688

A:Accession: AG3572

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-537 <KUR>

A:Cross-references: GB:AE008918; PIDN:AAL53746.1; PID:G17984672; GSPDB:GN00191

A:Experimental source: strain 16M

C:Genetics:

A:Gene: BMEI10504

A:Map position: II

Query Match 9.5%; Score 7; DB 2; Length 537;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
|||||

DB 172 PALTAVL 178
|||||

RESULT 11

S62421

endopeptidase La homolog (EC 3.4.21.-) PIM1 precursor, mitochondrial - fission yeast (SC

N:Alternate names: ATP-dependent proteinase lon homolog; hypothetical protein SPAC228F3.0

N:Contains: adenosinetriphosphatase (EC 3.6.1.3)

C:Species: Schizosaccharomyces pombe

C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 02-Jun-2003

C:Accession: S62421; T36185

R;Lye, G.; Churcher, C.M.
submitted to the EMBL Data Library, October 1995
A;Reference number: S62416
A;Accession: S62421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1067 <LYE>
A;Cross-references: EMBL:254285; NID:g1008429; PID:g1008434
R;Lye, G.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, October 1995
A;Reference number: Z21776
A;Accession: T38185
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1067 <LY2>
A;Cross-references: EMBL:254285; PID:CAA91071.1; GSPDB:GN000066; SPDB:SPAC22F3.06C
A;Experimental source: strain 972h; cosmid c22F3
C;Comment: This enzyme can also serve as a molecular chaperone and is essential for resp
C;Genetics:
A;Gene: SPDB:SPAC22F3.06C
A;Map position: 1L
C;Function:
A;Description: serine proteinase
A;Note: may be involved in heat shock response
C;Superfamily: ATP-dependent Lon protease
C;Keywords: ATP; DNA binding; heat shock; hydrolase; mitochondrial matrix; mitochondrion
F;578-585/Region: nucleotide-binding motif A (P-loop)
F;641-646/Region: nucleotide-binding motif B
F;584/Binding site: ATP (lys) #status predicted
F;946/Active site: Ser #status predicted

Query Match 9.5%; Score 7; DB 1; Length 1067;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
DB 244 PALTAVL 250

RESULT 12
FKBP-rapamycin-associated protein (FRAP) - human
C;Species: Homo sapiens (man)
C;Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 24-Nov-1999
C;Accession: S45340
R;Brown, E.J.; Albers, M.W.; Shin, T.B.; Ichikawa, K.; Keith, C.T.; Lane, W.S.; Schreiber
Nature 369, 756-758, 1994
A;Title: A mammalian protein targeted by G1-arresting rapamycin-receptor complex.
A;Reference number: S45340; MUID:94277209; PMID:8008069
A;Accession: S45340
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2549 <BRO>
A;Cross-references: GB:L34075; NID:g508481; PID:AAA58486.1; PID:g508482
C;Genetics:
A;Gene: GDB:FRAP1; FRAP; RAFT1
A;Cross-references: GDB:597698; OMIM:601231
A;Map position: lp36-lp36
C;Superfamily: yeast TOR2 protein

Query Match 9.5%; Score 7; DB 2; Length 2549;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
DB 515 PALTAVL 521

RESULT 13
A54837
rapamycin/FKBP12 target 1 - rat

N;Alternate names: RAFT
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Apr-1995 #sequence_revision 12-Apr-1995 #text_change 24-Nov-1999
C;Accession: A54837
R;Sabatini, D.M.; Erdjument-Bromage, H.; Lui, M.; Tempst, P.; Snyder, S.H.
Cell 78, 35-43, 1994
A;Title: RAFT1: a mammalian protein that binds to FKBP12 in a rapamycin-dependent fashion.
A;Reference number: A54837; MUID:94306515; PMID:7518356
A;Accession: A54837
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2549 <SAB>
A;Cross-references: GB:U11681; NID:g511228; PID:AAA20091.1; PID:g511229
C;Superfamily: yeast TOR2 protein

Query Match 9.5%; Score 7; DB 2; Length 2549;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
DB 515 PALTAVL 521

RESULT 14
CB1686
hypothetical protein TC0601 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: CB1686
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: CB1686
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-60 <TET>
A;Cross-references: GB:AE002328; GB:AE002160; NID:g7190627; PIDN:AAF39433.1; PID:g71906;
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0601

Query Match 8.1%; Score 6; DB 2; Length 60;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 49 SLWLVG 54
DB 17 SLWLVG 22

RESULT 15
D75413
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C;Accession: D75413
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma,
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: D75413
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-70 <WH1>
A;Cross-references: GB:AE001976; GB:AE000513; NID:g6459034; PIDN:AAF10865.1; PID:g645904
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1286

A;Map position: 1

Query Match 8.1%; Score 6; DB 2; Length 70;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 46 LRCSLW 51
| | | | |
Db 46 LRCSLW 51

Search completed: March 14, 2004, 21:55:35
Job time : 9.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:52:59 ; Search time 11.2276 Seconds
(without alignments)
1391.692 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 74

Sequence: 1 ARAYXKMFMEKAGKWC.....ARLXPCGKVGMDVRRWS 74

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 809742 seqs, 21113259 residues

Word size : 0

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubaa/PTCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubaa/PTCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	97.3	72	74	13	US-10-057-510-2
2	97.3	72	74	13	US-10-057-510-2
3	97.3	72	74	13	US-10-057-510-2
4	97.3	72	74	13	US-10-057-510-2
5	97.3	72	74	13	US-10-057-510-2
6	97.3	72	74	13	US-10-057-510-2
7	97.3	72	74	13	US-10-057-510-2
8	97.3	72	74	13	US-10-057-510-2
9	97.3	72	74	13	US-10-057-510-2
10	97.3	72	74	13	US-10-057-510-2
11	97.3	72	74	13	US-10-057-510-2
12	97.3	72	74	13	US-10-057-510-2
13	97.3	72	74	13	US-10-057-510-2
14	97.3	72	74	13	US-10-057-510-2
15	97.3	72	74	13	US-10-057-510-2

Sequence 8, Appli
Sequence 2869, Ap
Sequence 336, App
Sequence 15, Appl
Sequence 7, Appl
Sequence 6, Appl
Sequence 13, Appl
Sequence 21, Appl
Sequence 46, Appl
Sequence 282, App
Sequence 282, App
Sequence 280, App
Sequence 281, App
Sequence 280, App
Sequence 281, App
Sequence 8, Appl
Sequence 2, Appl
Sequence 10, Appl
Sequence 250, App
Sequence 16, Appl
Sequence 4777, Ap
Sequence 332, App
Sequence 3, Appl
Sequence 36, Appl
Sequence 8, Appl
Sequence 28, Appl
Sequence 5, Appl
Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-10-057-510-2

Sequence 2, Application US/10057510
Publication No. US20020098580A1
GENERAL INFORMATION:
APPLICANT: Nandabalan, Krishnan
APPLICANT: Yang, Meijia
APPLICANT: Schulz, Vincent
APPLICANT: Curagen Corporation
TITLE OF INVENTION: MDI INTERACTING PROTEIN AND METHODS OF USE THEREOF
FILE REFERENCE: 15966-524 MDM US
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: USSN 09/510,252
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: USSN 60/121,192
PRIOR FILING DATE: 1999-02-23
PRIOR APPLICATION NUMBER: USSN 60/122,643
PRIOR FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 74
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Any X can be any amino acid.
US-10-057-510-2

Query Match 97.3%; Score 72; DB 13; Length 74;
Best Local Similarity 100.0%; Pred. No. 5e-66; 0; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 0;

Qy 1 ARAYXKMFMEKAGKWCCKMKLIIDTFPFIIVAPALTAVLSQRLCSLWLVGARLXPC 60
Db 1 ARAYXKMFMEKAGKWCCKMKLIIDTFPFIIVAPALTAVLSQRLCSLWLVGARLXPC 60
Qy 61 GKVEGMDVRRWS 74

```
Db      61  GKVEGMDVWRRWS 74
|||||
Db      108 VEGMDVW 114
|||||

RESULT 4
US-10-369-493-23099
; Sequence 23099, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23099
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-369-493-23099

Query Match          9.5%; Score 7; DB 15; Length 469;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      38  LTAVLSC 44
|||||
Db      291  LTAVLSC 297

RESULT 5
US-09-738-626-4343
; Sequence 4343, Application US/09738626
; Publication No. US20020197603A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4343
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4343

Query Match          9.5%; Score 7; DB 9; Length 568;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db      63  GKVEGMDVWRRWS 74
|||||
Db      108 VEGMDVW 114
|||||

RESULT 2
US-10-156-761-12318
; Sequence 12318, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 12318
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-12318

Query Match          9.5%; Score 7; DB 14; Length 200;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      36  PALTAVL 42
|||||
Db      54  PALTAVL 60

RESULT 3
US-09-880-748-1294
; Sequence 1294, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; CURRENT FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1294
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1294

Query Match          9.5%; Score 7; DB 10; Length 247;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      63  VEGMDVW 69
|||||
```



```

; SEQ ID NO 2079
; LENGTH: 1067
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2079

Query Match          9.5%; Score 7; DB 15; Length 1067;
Best Local Similarity 100.0%; Pred No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      36 PALTAVL 42
      |||||
Db      244 PALTAVL 250

RESULT 8
US-09-950-634-3
; Sequence 3, Application US/09950634
; Publication No. US20030032775A1
; GENERAL INFORMATION:
; APPLICANT: Molnar-Kimber, Katherine L.
; Failli, Amedeo F.
; Caggiano, Thomas J.
; Nakanishi, Koji
; Chen, Yangdiu
; TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/950,634
; FILING DATE: 13-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/471,112
; FILING DATE: 06-JUN-1995
; APPLICATION NUMBER: US 08/384,524
; FILING DATE: 13-FEB-1995
; APPLICATION NUMBER: US 08/312,023
; FILING DATE: 26-SEP-1995
; APPLICATION NUMBER: US 08/207,975
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Siekman, Michael T.
; REGISTRATION NUMBER: 36,276
; REFERENCE/DOCKET NUMBER: 01142.0058-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2549 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-950-634-3

Query Match          9.5%; Score 7; DB 10; Length 2549;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 36 PALTAVL 42
Db 515 PALTAVL 521

RESULT 9
US-09-839-447A-87
; Sequence 87, Application US/09839447A
; Patent No. US20020058247A1
; GENERAL INFORMATION:
; APPLICANT: Sallberg, Matti
; TITLE OF INVENTION: SYNTHETIC PEPTIDES THAT BIND TO THE
; FILE REFERENCE: HEPATITIS B VIRUS CORE AND E ANTIGENS
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US/09/839,447A
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/556605
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial Peptide
US-09-839-447A-87

Query Match 8.1%; Score 6; DB 9; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAV 41
Db 3 PALTAV 8

RESULT 10
US-10-369-060A-87
; Sequence 87, Application US/10369060A
; Publication No. US20030235815A1
; GENERAL INFORMATION:
; APPLICANT: Sallberg, Matti
; TITLE OF INVENTION: SYNTHETIC PEPTIDES THAT BIND TO THE
; FILE REFERENCE: HEPATITIS B VIRUS CORE AND E ANTIGENS
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 09/839,447
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 09/556,605
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 87
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mus musculus
US-10-369-060A-87

Query Match 8.1%; Score 6; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAV 41
Db 3 PALTAV 8

US-10-608-541-51
; Sequence 51, Application US/10608541
; Publication No. US20040019189A1
; GENERAL INFORMATION:
; APPLICANT: Matti Sallberg
; TITLE OF INVENTION: LIGAND/RECEPTOR SPECIFICITY EXCHANGERS
; FILE REFERENCE: TRIPEP.007CF3C1
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: US/10/608,541
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/664,945
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 09/532,106
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 08/737,085
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: PCT/SE 95/00468
; PRIOR FILING DATE: 1995-04-27
; PRIOR APPLICATION NUMBER: SE 9401460
; PRIOR FILING DATE: 1994-04-28
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 51
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antigenic domain peptide
US-10-608-541-51

Query Match 8.1%; Score 6; DB 15; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.1e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAV 41
Db 3 PALTAV 8

RESULT 12
US-09-839-447A-84
; Sequence 84, Application US/09839447A
; Patent No. US20020058247A1
; GENERAL INFORMATION:
; APPLICANT: Sallberg, Matti
; TITLE OF INVENTION: SYNTHETIC PEPTIDES THAT BIND TO THE
; FILE REFERENCE: TRIPEP.020CF1
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US/09/839,447A
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial Peptide
US-09-839-447A-84

Query Match 8.1%; Score 6; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAV 41
Db 1 PALTAV 6

```

RESULT 13
US-10-369-060A-84
; Sequence 84, Application US/10369060A
; Publication No. US20030235815A1
; GENERAL INFORMATION:
; APPLICANT: Saliberg, Matti
; TITLE OF INVENTION: SYNTHETIC PEPTIDES THAT BIND TO THE
; TITLE OF INVENTION: HEPATITIS B VIRUS CORE AND E ANTIGENS
; FILE REFERENCE: TRIPEP.020CP1C1
; CURRENT APPLICATION NUMBER: US/10/369,060A
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: 09/839,447
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 09/555,605
; PRIOR FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mus musculus
US-10-369-060A-84

Query Match      8.1%; Score 6; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      36 PALTAV 41
DB      1 PALTAV 6

RESULT 14
US-10-608-541-48
; Sequence 48, Application US/10608541
; Publication No. US20040019189A1
; GENERAL INFORMATION:
; APPLICANT: Matti Saliberg
; TITLE OF INVENTION: LIGAND/RECEPTOR SPECIFICITY EXCHANGERS
; TITLE OF INVENTION: THAT REDIRECT ANTIBODIES TO RECEPTORS ON A PATHOGEN
; FILE REFERENCE: TRIPEP.007CP3C1
; CURRENT APPLICATION NUMBER: US/10/608,541
; CURRENT FILING DATE: 2003-06-27
; PRIOR APPLICATION NUMBER: 09/664,945
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/532,106
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 09/246,258
; PRIOR FILING DATE: 1999-02-08
; PRIOR APPLICATION NUMBER: 09/737,085
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: PCT/SE 95/00468
; PRIOR FILING DATE: 1995-04-27
; PRIOR APPLICATION NUMBER: SE 9401460
; PRIOR FILING DATE: 1994-04-28
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Antigenic domain peptide
US-10-608-541-48

Query Match      8.1%; Score 6; DB 15; Length 11;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      36 PALTAV 41
DB      1 PALTAV 6

```

```

DB      1 PALTAV 6

RESULT 15
US-09-839-666-8
; Sequence 8, Application US/09839666
; Patent No. US20020025513A1
; GENERAL INFORMATION:
; APPLICANT: SALLBERG, MATTI
; TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
; EXCHANGER
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DARBY & DARBY PC
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/839,666
; FILING DATE: 19-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/737,085
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 3846/0C569
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-527-7659
; TELEFAX: 212-753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-839-666-8

Query Match      8.1%; Score 6; DB 9; Length 12;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      36 PALTAV 41
DB      4 PALTAV 9

Search completed: March 14, 2004, 21:57:31
Job time : 12.2276 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:51:29 ; Search time 7.91034 Seconds
(without alignments)
482.953 Million cell updates/sec

Title: US-10-057-510-2

Perfect score: 74

Sequence: 1 ARAYXKXFMFKAGKWC.....ARLXPCGKVGMDVRRWS 74

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pbp.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pbp.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pbp.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pbp.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pbp.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pbp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	72	97.3	74	US-09-510-252-2	Sequence 2, Appli
2	7	9.5	2549	US-08-471-112A-3	Sequence 3, Appli
3	7	9.5	2549	US-08-265-967C-1	Sequence 1, Appli
4	7	9.5	2549	US-08-305-790B-2	Sequence 2, Appli
5	7	9.5	2549	PCT-US95-06722-12	Sequence 12, Appli
6	6	8.1	7	US-09-147-933-11	Sequence 11, Appli
7	6	8.1	9	US-09-664-945-51	Sequence 51, Appli
8	6	8.1	11	US-09-664-945-48	Sequence 48, Appli
9	6	8.1	12	US-08-737-085A-8	Sequence 8, Appli
10	6	8.1	12	US-09-248-258-8	Sequence 8, Appli
11	6	8.1	12	US-09-532-106-8	Sequence 8, Appli
12	6	8.1	12	US-09-839-686-8	Sequence 8, Appli
13	6	8.1	14	US-09-701-432-11	Sequence 11, Appli
14	6	8.1	17	US-08-836-922-18	Sequence 18, Appli
15	6	8.1	17	US-08-602-999A-336	Sequence 336, App
16	6	8.1	17	US-09-500-124-336	Sequence 336, App
17	6	8.1	17	US-09-639-681-18	Sequence 18, Appli
18	6	8.1	21	US-08-878-277A-13	Sequence 13, Appli
19	6	8.1	26	US-08-737-085A-15	Sequence 15, Appli
20	6	8.1	26	US-09-246-258-15	Sequence 15, Appli
21	6	8.1	26	US-09-532-106-15	Sequence 15, Appli
22	6	8.1	26	US-09-839-666-15	Sequence 15, Appli
23	6	8.1	49	US-10-037-927B-40	Sequence 40, Appli
24	6	8.1	79	US-09-621-976-7497	Sequence 7497, Ap
25	6	8.1	91	US-09-489-039A-10828	Sequence 10828, A
26	6	8.1	121	US-09-252-991A-27340	Sequence 27340, A
27	6	8.1	125	US-09-252-991A-32594	Sequence 32594, A

28 6 8.1 137 4 US-09-252-991A-29615. Sequence 29615, A
29 6 8.1 139 3 US-09-289-349-10. Sequence 10, Appl
30 6 8.1 141 4 US-09-107-532A-4395. Sequence 4395, Ap
31 6 8.1 147 4 US-09-621-976-4554. Sequence 4554, AD
32 6 8.1 158 6 5312912-2. Patent No. 5312912
33 6 8.1 167 4 US-09-543-681A-7113. Sequence 7113, Ap
34 6 8.1 169 1 US-08-145-995A-7. Sequence 7, Appli
35 6 8.1 169 2 US-08-451-747-7. Sequence 7, Appli
36 6 8.1 169 3 US-09-134-853-7. Sequence 7, Appli
37 6 8.1 171 1 US-08-145-995A-10. Sequence 10, Appl
38 6 8.1 171 2 US-08-451-747-10. Sequence 10, Appl
39 6 8.1 171 3 US-09-134-852-10. Sequence 10, Appl
40 6 8.1 173 4 US-09-328-352-7080. Sequence 7080, Ap
41 6 8.1 178 4 US-09-328-352-4981. Sequence 4981, Ap
42 6 8.1 184 4 US-09-565-423-11. Sequence 11, Appl
43 6 8.1 194 4 US-09-489-039A-12423. Sequence 12423, A
44 6 8.1 217 4 US-09-489-039A-13393. Sequence 13393, A
45 6 8.1 235 4 US-09-252-991A-30499. Sequence 30499, A

ALIGNMENTS

RESULT 1

US-09-510-252-2
; Sequence 2, Application US/09510252
; Patent No. 6372490
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meijia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDM INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT APPLICATION NUMBER: US/09/510,252
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USSN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USSN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Any X can be any amino acid.
US-09-510-252-2

Query Match 97.3%; Score 72; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.2e-70; Indels 0; Gaps 0;
Matches 74; Conservative 0; Mismatches 0

Qy 1 ARAYXKXFMFKAGKWCMPKLIIDTPFISVAPALTAIVLSQRLCSLWLVGARLXPC 60

Db 1 ARAYXKXFMFKAGKWCMPKLIIDTPFISVAPALTAIVLSQRLCSLWLVGARLXPC 60

Qy 61 GKVEGMDVRRWS 74

Db 61 GKVEGMDVRRWS 74

RESULT 2

US-08-471-112A-3
; Sequence 3, Application US/08471112A
; Patent No. 6313264
; GENERAL INFORMATION:
; APPLICANT: Molnar-Kimber, Katherine L.
; APPLICANT: Failli, Amedeo F.
; APPLICANT: Caggiano, Thomas J.
; APPLICANT: Nakanishi, Koji
; APPLICANT: Chen, Yanqiu

```

; TITLE OF INVENTION: EFFECTOR PROTEINS OF RAPAMYCIN
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,112A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/384,524
; FILING DATE: 13-FEB-1995
; PRIOR APPLICATION DATA: US 08/312,023
; APPLICATION NUMBER: US 08/312,023
; FILING DATE: 26-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/207,975
; FILING DATE: 08-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Siekman, Michael T.
; REGISTRATION NUMBER: 36,276
; REFERENCE/DOCKET NUMBER: 01142.0058-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 3:
; LENGTH: 2549 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-471-112A-3

Query Match 9.5%; Score 7; DB 4; Length 2549;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
Db 515 PALTAVL 521

RESULT 3
US-08-265-967C-1
; Sequence 1, Application US/08265967C
; Patent No. 6476200
; GENERAL INFORMATION:
; APPLICANT: SABATINI, DAVID M.
; APPLICANT: ERDJUMENT-BROMAGE, HEDIYE
; APPLICANT: LUI, MARY
; APPLICANT: TEMPEST, PAUL
; APPLICANT: SNYDER, SOLOMON H.
; TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & ALLEGRETTI, LTD
; STREET: 1001 G STREET, N.W., 11TH FLOOR
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001-4597

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/265,967C
; FILING DATE: 27-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.46363
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2549 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Rattus rattus
; TISSUE TYPE: pheochromocytoma
; CELL TYPE: PC12
; US-08-265-967C-1

Query Match 9.5%; Score 7; DB 4; Length 2549;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
Db 515 PALTAVL 521

RESULT 4
US-08-305-790B-2
; Sequence 2, Application US/08305790B
; Patent No. 6492106
; GENERAL INFORMATION:
; APPLICANT: SABATINI, DAVID M.
; APPLICANT: ERDJUMENT-BROMAGE, HEDIYE
; APPLICANT: LUI, MARY
; APPLICANT: TEMPEST, PAUL
; APPLICANT: SNYDER, SOLOMON H.
; TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & ALLEGRETTI, LTD
; STREET: 1001 G STREET, N.W., 11TH FLOOR
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/305,790B
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/265,967
; FILING DATE: 27-JUN-1994
; ATTORNEY/AGENT INFORMATION:

```

NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.47225
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BEME UT
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2549 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Rattus rattus
TISSUE TYPE: pheochromocytoma
CELL TYPE: PC12
US-08-305-790B-2

Query Match 9.5%; Score 7; DB 4; Length 2549;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
Db 515 PALTAVL 521

RESULT 5

PCT-US95-06722-12
Sequence 12, Application PC/TUS9506722
GENERAL INFORMATION:

APPLICANT: Immunosuppressant Target Proteins
NUMBER OF SEQUENCES: 25
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII (text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/06722
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 27-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/250,795
FILING DATE: 20-DEC-1994
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 2549 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-06722-12

Query Match 9.5%; Score 7; DB 5; Length 2549;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
Db 515 PALTAVL 521

RESULT 6

US-09-147-933-11
Sequence 11, Application US/09147933A
Patent No. 6168917

GENERAL INFORMATION:
APPLICANT: Kilpatrick, David
TITLE OF INVENTION: DETECTION AND IDENTIFICATION OF
TELECOMMUNICATION INFORMATION:
FILE REFERENCE: 62242/US
CURRENT APPLICATION NUMBER: US/09/147,933A
CURRENT FILING DATE: 1999-07-09
EARLIER APPLICATION NUMBER: PCT/US97/17734
EARLIER FILING DATE: 1997-10-01
EARLIER APPLICATION NUMBER: U. S. 60/027,353
EARLIER FILING DATE: 1996-10-02
NUMBER OF SEQ ID NOS: 96
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: peptide
US-09-147-933-11

Query Match 8.1%; Score 6; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAV 41
Db 1 PALTAV 6

RESULT 7

US-09-664-945-51
Sequence 51, Application US/09664945
Patent No. 6660842
GENERAL INFORMATION:
APPLICANT: Matti Sallberg
TITLE OF INVENTION: LIGAND/RECEPTOR SPECIFICITY EXCHANGERS
THAT REDIRECT ANTIBODIES TO RECEPTORS ON A PATHOGEN
FILE REFERENCE: TRIPEP.007CP3
CURRENT APPLICATION NUMBER: US/09/664,945
CURRENT FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 09/532,106
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 09/246,258
PRIOR FILING DATE: 1999-02-08
PRIOR APPLICATION NUMBER: 08/737,085
PRIOR FILING DATE: 1996-12-27
PRIOR APPLICATION NUMBER: SE 9401460
PRIOR FILING DATE: 1994-04-28
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 51
LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antigenic domain peptide
US-09-664-945-51

Query Match 8.1%; Score 6; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAV 41
Db 3 PALTAV 8

RESULT 8

US-09-664-945-48
Sequence 48, Application US/09664945
Patent No. 6660842
GENERAL INFORMATION:

APPLICANT: Matti Sallberg
TITLE OF INVENTION: LIGAND/RECEPTOR SPECIFICITY EXCHANGERS
FILE REFERENCE: TRIPEP.007CP3
CURRENT APPLICATION NUMBER: US/09/664,945
PRIOR FILING DATE: 2000-09-19
PRIOR APPLICATION NUMBER: 09/532,106
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 09/246,258
PRIOR FILING DATE: 1999-02-08
PRIOR APPLICATION NUMBER: 08/737,085
PRIOR FILING DATE: 1996-12-27
PRIOR APPLICATION NUMBER: SE 9401460
PRIOR FILING DATE: 1994-04-28
NUMBER OF SEQ ID NOS: 105
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 48
LENGTH: 11
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Antigenic domain peptide
US-09-664-945-48

Query Match 8.1%; Score 6; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAV 41
Db 1 PALTAV 6

RESULT 9
US-08-737-085A-8
Sequence 8, Application US/08737085A
Patent No. 5869232
GENERAL INFORMATION:
APPLICANT: SALLBERG, MATTI
TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
EXCHANGER
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: DARBY & DARBY PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,085A
FILING DATE: 27-DEC-1996
CLASSIFICATION: 426
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 3846/0C569
TELEPHONE: 212-527-7659
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-737-085A-8

Query Match 8.1%; Score 6; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAV 41
Db 4 PALTAV 9

RESULT 10
US-09-246-258-8
Sequence 8, Application US/09246258
Patent No. 6040137
GENERAL INFORMATION:
APPLICANT: SALLBERG, MATTI
TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY
EXCHANGER
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: DARBY & DARBY PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/246,258
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/737,085
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 3846/0C569
TELEPHONE: 212-527-7659
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-246-258-8

Query Match 8.1%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAV 41
Db 4 PALTAV 9

RESULT 11
US-09-532-106-8
Sequence 8, Application US/09532106
Patent No. 6245895
GENERAL INFORMATION:

APPLICANT: SALLBERG, MATTI
TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY EXCHANGER
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: DARBY & DARBY PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/532,106
FILING DATE: 21-Mar-2000
Prior APPLICATION DATA:
APPLICATION NUMBER: US/08/737,085A
FILING DATE: 27-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 3846/OC569
TELEPHONE: 212-527-7659
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-532-106-8
Query Match 8.1%; Score 6; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 36 PALTAV 41
DB 4 PALTAV 9
RESULT 12
US-09-839-666-8
Sequence 8, Application US/09839666
Patent No. 6469143
GENERAL INFORMATION:
APPLICANT: SALLBERG, MATTI
TITLE OF INVENTION: ANTIGEN/ANTIBODY SPECIFICITY EXCHANGER
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: DARBY & DARBY PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/839,666
FILING DATE: 19-Apr-2001

CLASSIFICATION: <Unknown>
Prior APPLICATION DATA:
APPLICATION NUMBER: 08/737,085
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Green, Reza
REGISTRATION NUMBER: 38,475
REFERENCE/DOCKET NUMBER: 3846/OC569
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7659
TELEFAX: 212-753-6237
TELEX: 236687
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-839-666-8
Query Match 8.1%; Score 6; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 36 PALTAV 41
DB 4 PALTAV 9
RESULT 13
US-09-701-432-11
Sequence 11, Application US/09701432
Patent No. 6669939
GENERAL INFORMATION:
APPLICANT: SCHATZ, OCTAVIAN
APPLICANT: HAAS, JURGEN
TITLE OF INVENTION: NOVEL (POLY)PEPTIDES WHICH REPRESENT THE EPITOPES
OF HUMAN HERPESVIRUS TYPE 8
FILE REFERENCE: 028622/0104
CURRENT APPLICATION NUMBER: US/09/701,432
CURRENT FILING DATE: 2001-02-08
Prior APPLICATION NUMBER: PCT/EP99/03719
Prior FILING DATE: 1999-05-28
Prior APPLICATION NUMBER: DE 198 24 244.1
Prior FILING DATE: 1998-05-29
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 14
TYPE: PRT
ORGANISM: Poliovirus
US-09-701-432-11
Query Match 8.1%; Score 6; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 7.3;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 36 PALTAV 41
DB 5 PALTAV 10
RESULT 14
US-08-836-922-18
Sequence 18, Application US/08836922
Patent No. 6159711
GENERAL INFORMATION:
APPLICANT: INNES PROUDFOOT, AMANDA ELIZABETH
APPLICANT: WELLS, TIMOTHY NIGEL CARL
TITLE OF INVENTION: RANTES PEPTIDE AND FRAGMENTS AND
COMPOSITIONS COMPRISING IT FOR TREATMENT OF INFLAMMATION

NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/836,922
FILING DATE: 23-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9424835.8
FILING DATE: 08-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9512319.6
FILING DATE: 16-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1430-163
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4011
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-836-922-18

Query Match 8.1%; Score 6; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 PKLIID 28
Db 12 PKLIID 17

RESULT 15

US-08-602-999A-336
Sequence 336, Application US/08602999A
Patent No. 6184205

GENERAL INFORMATION:

APPLICANT: SPARKS, Andrew B.
APPLICANT: KAY, Brian K.
APPLICANT: THORN, Judith M.
APPLICANT: QUILLIAM, Lawrence A.
APPLICANT: DER, Channing J.
APPLICANT: FOWLKES, Dana M.
APPLICANT: RIDER, James E.
TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
TITLE OF INVENTION: ISOLATING AND USING SAME
NUMBER OF SEQUENCES: 467
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/602,999A
FILING DATE: 16-FEB-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-202
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 336:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-602-999A-336

Query Match 8.1%; Score 6; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 CWCXMP 23
Db 2 CWCXMP 7

Search completed: March 14, 2004, 21:56:26
Job time : 8.91035 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:45:48 ; Search time 18.1172 Seconds
(without alignments)
1154.067 Million cell updates/sec

Title: US-10-057-510-2
 Perfect score: 74
 Sequence: 1 ARAYXKMFSEMFNEKAGKCWC.....ARLYPCGKVEGMDVWRRRUS 74

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 45 summaries

```
Database : A_Geneseq_29Jan04.*
1: geneseq1990s.*
2: geneseq1990s.*
3: geneseq2000s.*
4: geneseq2001s.*
5: geneseq2002s.*
6: geneseq2003as.*
7: geneseq2003bs.*
8: geneseq2004s.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	72	97.3	74	3	AA080845	Aab080845 A human M
2	7	9.5	247	5	ABP45283	Abp45283 Human Bly
3	7	9.5	261	5	AB548347	Abp48347 Listeria
4	7	9.5	434	6	ABU34740	Abu34740 Protein e
5	7	9.5	434	6	ABU37031	Abu37031 Protein e
6	7	9.5	447	6	ABU18962	Abu18962 Protein e
7	7	9.5	451	6	ABU33736	Abu33736 Protein e
8	7	9.5	464	6	ABU18097	Abu18097 Protein e
9	7	9.5	547	7	AD881207	Ad881207 Orf17, SE
10	7	9.5	568	4	AA879753	AA879753 Corynebac
11	7	9.5	568	4	AA880200	AA880200 Corynebac
12	7	9.5	568	4	AAU71921	AAU71921 C. glutam
13	7	9.5	568	4	AAQ90589	AAQ90589 C. glutami
14	7	9.5	647	4	AB859908	AB859908 Drosophil
15	7	9.5	1417	7	AD562172	Ad562172 Human Pro
16	7	9.5	2549	2	AA881730	AA881730 Sirolimus
17	7	9.5	2549	2	AAW04235	AAW04235 Human RAP
18	7	9.5	2549	2	AAW56027	AAW56027 FRAP (cor
19	7	9.5	2549	5	AAE13359	AAe13359 PUC19-Sep
20	7	9.5	2549	6	ABU08621	Abu08621 Rat (rapa
21	7	9.5	2549	6	ABU61979	Abu61979 Human sir
22	7	9.5	2549	7	AD862170	Ad862170 Rat Prote
23	7	9.5	6797	7	AA831358	AA831358 Pimaricin
24	6	8.1	6	8	AAW59308	AAW59308 Non-polio
25	6	8.1	7	2	AAV50068	AAV50068 Coxsackie

RESULT 1
AAB08845
ID AAB08845 standard; protein, 74 AA.
XX
XX AAB08845;
XX AC
XX AC
XX DT
XX C2-JAN-2001 (first entry)
XX
XX A human MDM2 interacting polypeptide (MDMIP).

XX	WO2000050590-A1.
PN	
XX	
XX	31-AUG-2000.
PD	
XX	
XX	23-FEB-2000; 2000WO-US004582.
PF	
XX	
XX	23-FEB-1999; 99US-0121192P.
PR	
PR	03-MAR-1999; 99US-0122643P.
PP	
PP	22-FEB-2000; 2000US-00510282.

XX (CURA-) CURAGEN CORP.
PA
XX
XX
XX
PI Nandabalan K, Yang M, Schulz VP;
XX
XX WPI; 2000-558398/51.
DR N-PSDB: AAA75041.
DR

Novel MDM2 interacting protein useful for treating or preventing disorders involving aberrant levels of MDM2 and/or MDM2-interacting proteins, comprises a specific amino acid sequence.

XX
ps
Claim 10: Fig 1: 78pp: English.

xx The present sequence represents a human MDM2 interacting polypeptide
cc (MDMip). MDMip was identified using a yeast two hybrid system, using a
cc

CC fragment of MD2 as the bait protein. The MD2 polypeptide is useful for
 CC detecting and removing MD2 polypeptides in a biological sample by
 CC forming MD2-MD2 complexes. MD2 and MD2 are useful to identify
 CC compounds or other agents which modulate the activity of MD2 and/or
 CC MD2-mediated processes. Agents that modulate the function of MD2/MD2
 CC complexes are useful for treating and preventing a disease or disorder
 CC involving aberrant levels of MD2 or MD2. MD2 is also useful for
 CC treating diseases caused by aberrant levels of expression of MD2 genes,
 CC such as disorders of cell cycle progression, cell differentiation, and
 CC transcriptional control, including cancers such as human sarcoma, glioma,
 CC squamous cell carcinoma, breast cancer, astrocytoma, leukemia and
 CC lymphoma, and tumorigenesis. MD2 and MD2 nucleic acids are useful in
 CC gene therapy
 CC
 CC Sequence 74 AA;

Query Match 97.3%; Score 72; DB 3; Length 74;
 Best Local Similarity 100.0%; Pred. No. 1.3e-67; Indels 0; Gaps 0;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ARAYXKWFNFMEKAGCKCKPKLIIDTPFIVAPALTAVLSQRLCSLWLVGARLXPC 60
 DB 1 ARAYXKWFNFMEKAGCKCKPKLIIDTPFIVAPALTAVLSQRLCSLWLVGARLXPC 60
 QY 61 GKVEGMDVWRRWS 74
 DB 61 GKVEGMDVWRRWS 74

RESULT 2

ABP45283

ID ABP45283 standard; protein; 247 AA.

AC ABP45283;

XX 19-AUG-2002 (first entry)

DE Human Blys binding scFv SEQ ID 1294.

XX Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;
 KW immunosuppressive; immunostimulant; antirheumatic;
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.

XX Homo sapiens.

OS WO200202641-A1.

PN 10-JAN-2002.

PD 15-JUN-2001; 2001WO-US019110.

PF 16-JUN-2000; 2000US-0212210P.

PR 17-OCT-2000; 2000US-0240816P.

PR 16-MAR-2001; 2001US-0276248P.

PR 21-MAR-2001; 2001US-0277379P.

PR 25-MAY-2001; 2001US-0293499P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

XX WPI; 2002-114799/15.

XX Antibodies against B lymphocyte stimulating polypeptides, useful for the
 PT diagnosis and treatment of cancers and immune disorders.

XX Claim 1; Page 1947-1948; 3148pp; English.

XX This invention describes novel antibodies that immunospecifically bind to

CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the
 CC tumour necrosis factor (TNF) super family and induces B cell
 CC proliferation and differentiation. The antibodies of the invention have
 CC cyclostatic, immunosuppressive, immunostimulant, immunomodulatory,
 CC antirheumatic and antiAIDS activity and can be used in vaccines to
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys
 CC and so may be used to detect and quantitate the presence of Blys in
 CC biological samples and may be used in this way to diagnose disease
 CC associated with aberrant expression of Blys. They may also be
 CC administered to treat diseases associated with aberrant Blys expression
 CC and activity such as cancer, immune, and autoimmune disorders and
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and
 CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent
 CC the antibodies and fragments of the antibodies described in the method of
 CC the invention

XX Sequence 247 AA;

Query Match 9.5%; Score 7; DB 5; Length 247;

Best Local Similarity 100.0%; Pred. No. 52;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 VEGMDVW 69

DB 108 VEGMDVW 114

RESULT 3

ABB48347

ID ABB48347 standard; protein; 261 AA.

XX ABB48347;

AC 05-FEB-2002 (first entry)

XX Listeria monocytogenes protein #1051.

DE Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
 KW vitamin B12; bacterial infection; disease.

XX Listeria monocytogenes.

OS WO200177335-A2.

PN 18-OCT-2001.

PF 11-APR-2001; 2001WO-FR001118.

PR 11-APR-2000; 2000FR-00004629.

XX (INSP) INST PASTEUR.

XX Buchrieser C, Prangeul L, Couve E, Rusniok C, Faihi H, Dehoux P;
 PI Dussurget O, Chetoui F, Nedjari H, Glaser P, Kunst F, Cossart P;
 PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;
 PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
 PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
 PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
 PI Madueno E, De Pablo B, Wehland J, Kaerst U, Entian K, Hauf J;
 PI Rose M, Voss H;
 XX WPI; 2002-010914/01.

XX Genomic sequence for Listeria monocytogenes, useful e.g. for treatment
 PT and prevention of Listeria and related bacterial infections, and related
 PT polypeptides.

XX Claim 6; SEQ ID NO 1052; 192pp; French.

XX The present invention relates to the genome sequence of Listeria
 CC monocytogenes EGD-e (see ABA03041). The genome sequence and fragments of
 CC it are useful for selecting probes and primers for detecting genes in L.

CC monocytes and related organisms, and for studying genetic
 CC polymorphisms and other genomes. The present sequence is a protein
 CC encoded by the genome sequence of the present invention. Proteins
 CC expressed from the genome sequence are useful for raising specific
 CC antibodies, identification of L. monocytogenes and related organisms, and
 CC for biosynthesis and biodegradation, especially biosynthesis of Vitamin
 CC B12. The genome sequence and proteins encoded by it are also useful for
 CC selecting compounds that regulate gene expression and cell replication
 CC and modulate L. monocytogenes-related diseases. In addition, the genome
 CC sequence and proteins encoded by it are useful in pharmaceutical and
 CC vaccine compositions for the treatment or prevention of infections by L.
 CC monocytogenes and related organisms. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 261 AA;

Query Match 9.5%; Score 7; DB 5; Length 261;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 MPKLIID 28
 |||||
 DB 1 MPKLIID 7

RESULT 4
 ABU34740
 ID ABU34740 standard; protein; 434 AA.

XX AC ABU34740;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #20267.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Mycobacterium bovis.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.

DR N-PSDB; ACA38610.

PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 62664; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 623 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway;
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than S. aureus, S. typhimurium,
 CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 434 AA;

Query Match 9.5%; Score 7; DB 6; Length 434;
 Best Local Similarity 100.0%; Pred. No. 84;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 LTAVLSC 44

|||||
 DB 271 LTAVLSC 277

RESULT 5
 ABU37031

ID ABU37031 standard; protein; 434 AA.

XX AC ABU37031;

DT 19-JUN-2003 (first entry)

DE Protein encoded by Prokaryotic essential gene #22558.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

OS Mycobacterium tuberculosis.

PN WO200277183-A2.

PD 03-OCT-2002.

PF 21-MAR-2002; 2002WO-US009107.

PR 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;

XX PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

DR WPI; 2003-029926/02.

DR N-PSDB; ACA40901.

PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 64955; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 434 AA;

XX

Query Match 9.5%; Score 7; DB 6; Length 434;

Best Local Similarity 100.0%; Pred. No. 84;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 LTAVLSC 44

Db 271 LTAVLSC 277

|||||

RESULT 6

ABU18962

ID ABU18962 standard; protein; 447 AA.

XX

AC ABU18962;

XX

XX 19-JUN-2003 (first entry)

DT

XX

XX Protein encoded by Prokaryotic essential gene #4489.

XX

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

XX Bacillus anthracis.

OS

XX

XX WO200277183-A2.

XX

XX 03-OCT-2002.

PD

XX

XX 21-MAR-2002; 2002WO-US009107.

PF

XX

XX 21-MAR-2001; 2001US-00815242.

PR

XX

XX 06-SEP-2001; 2001US-00948993.

PR

XX

XX 25-OCT-2001; 2001US-0342323P.

PR

XX

XX 08-FEB-2002; 2002US-00072851.

PR

XX

XX 06-MAR-2002; 2002US-0362699P.

PR

XX

XX

PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malore C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX

XX WPI; 2003-029926/02.

DR N-PSDB; ACA22832.

XX

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX

XX Claim 25; SEQ ID NO 46886; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of

CC the 6213 antisense sequences given in the specification where expression

CC of the nucleic acid inhibits proliferation of a cell. Also included are:

CC (1) a vector comprising a promoter operably linked to the nucleic acid

CC encoding a polypeptide whose expression is inhibited by the antisense

CC nucleic acid; (2) a host cell containing the vector; (3) an isolated

CC polypeptide or its fragment whose expression is inhibited by the

CC antisense nucleic acid; (4) an antibody capable of specifically binding

CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular

CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of

CC the gene product or that has an activity against a biological pathway

CC required for proliferation, or that inhibits cellular proliferation; (8)

CC identifying a gene required for cellular proliferation or the biological

CC pathway in which a proliferation-required gene or its gene product lies

CC or a gene on which the test compound that inhibits proliferation of an

CC organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene

CC product is overexpressed or underexpressed; (12) determining the extent

CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the

CC proliferation of an organism. The antisense nucleic acids are useful for

CC identifying proteins or screening for homologous nucleic acids required

CC for cellular proliferation to isolate candidate molecules for rational

CC drug discovery programs, or for screening homologous nucleic acids

CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,

CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of

CC the target prokaryotic essential genes. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 447 AA;

XX

Query Match 9.5%; Score 7; DB 6; Length 447;

Best Local Similarity 100.0%; Pred. No. 86;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 38 LTAVLSC 44

Db 286 LTAVLSC 292

|||||

RESULT 7

ABU33736

ID ABU33736 standard; protein; 451 AA.

XX

AC ABU33736;

XX

XX 19-JUN-2003 (first entry)

DT

XX

XX Protein encoded by Prokaryotic essential gene #19263.

XX

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

XX Mycobacterium avium.

OS

XX

XX WO200277183-A2.

PN

XX

Db 288 LTAVLC 294
|||||

RESULT 9

AD881207
ID ADE81207 standard; protein; 547 AA.

XX ADE81207;
AC

XX 29-JAN-2004 (first entry)
DT

DE Orf17, SEQ ID 35.
DE

XX KW ML-236B; HMG-CoA reducing enzyme; Orf17.
KW

XX Penicillium citrinum.
OS

XX JP2003116567-A.
PN

XX 22-APR-2003.
PD

XX PF 15-OCT-2001; 2001JP-00316578.
PF

XX 15-OCT-2001; 2001JP-00316578.
PR

XX (SANY) SANKYO CO LTD.
PA

XX WPI; 2003-817677/77.
DR

XX N-PSDB; ADE81206.
DR

XX Novel DNA associated with synthesis of ML-236B, useful for improving ML-236B production in ML-236B producing microbe.
PT

XX Example 8; SEQ ID NO 35; 142pp; Japanese.
PS

XX The present invention relates to a DNA sequence (I, ADE81173), which is associated with ML-236B synthesis. (I) is useful for improving ML-236B production in a HMG-CoA reducing-enzyme-inhibitor ML-236B producing microbe. The present sequence was used to illustrate the invention.
CC

XX Sequence 547 AA;
SQ

Query Match 9.5%; Score 7; DB 7; Length 547;
Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 PALTAVL 42
|||

Db 487 PALTAVL 493
|||||

RESULT 10

AB879753
ID AAB79753 standard; protein; 568 AA.

XX AAB79753;
AC

XX 30-APR-2001 (first entry)
DT

XX Corynebacterium glutamicum MP protein sequence SEQ ID NO:240.
DE

XX Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.

XX Corynebacterium glutamicum.
OS

XX WO200100843-A2.
PN

XX 04-JAN-2001.
PD

XX 23-JUN-2000; 2000WO-IB000923.
PF

XX 25-JUN-1999; 99US-0141031P.
PR

XX 01-JUL-1999; 99DE-01030476.
PR

XX 02-JUL-1999; 99US-0142101P.
PR

XX 08-JUL-1999; 99DE-01031415.
PR

XX 08-JUL-1999; 99DE-01031418.
PR

XX 08-JUL-1999; 99DE-01031419.
PR

XX 08-JUL-1999; 99DE-01031420.
PR

XX 08-JUL-1999; 99DE-01031424.
PR

XX 08-JUL-1999; 99DE-01031428.
PR

XX 08-JUL-1999; 99DE-01031434.
PR

XX 08-JUL-1999; 99DE-01031435.
PR

XX 08-JUL-1999; 99DE-01031443.
PR

XX 08-JUL-1999; 99DE-01031453.
PR

XX 08-JUL-1999; 99DE-01031457.
PR

XX 08-JUL-1999; 99DE-01031465.
PR

XX 08-JUL-1999; 99DE-01031478.
PR

XX 08-JUL-1999; 99DE-01031510.
PR

XX 08-JUL-1999; 99DE-01031541.
PR

XX 08-JUL-1999; 99DE-01031573.
PR

XX 08-JUL-1999; 99DE-01031592.
PR

XX 08-JUL-1999; 99DE-01031632.
PR

XX 08-JUL-1999; 99DE-01031634.
PR

XX 08-JUL-1999; 99DE-01031636.
PR

XX 09-JUL-1999; 99DE-01032125.
PR

XX 09-JUL-1999; 99DE-01032126.
PR

XX 09-JUL-1999; 99DE-01032130.
PR

XX 09-JUL-1999; 99DE-01032186.
PR

XX 09-JUL-1999; 99DE-01032206.
PR

XX 09-JUL-1999; 99DE-01032227.
PR

XX 09-JUL-1999; 99DE-01032228.
PR

XX 09-JUL-1999; 99DE-01032229.
PR

XX 09-JUL-1999; 99DE-01032230.
PR

XX 14-JUL-1999; 99DE-01032922.
PR

XX 14-JUL-1999; 99DE-01032926.
PR

XX 14-JUL-1999; 99DE-01032928.
PR

XX 14-JUL-1999; 99DE-01033004.
PR

XX 14-JUL-1999; 99DE-01033005.
PR

XX 14-JUL-1999; 99DE-01033006.
PR

XX 12-AUG-1999; 99US-0148613P.
PR

XX 27-AUG-1999; 99DE-01040764.
PR

XX 27-AUG-1999; 99DE-01040765.
PR

XX 27-AUG-1999; 99DE-01040766.
PR

XX 27-AUG-1999; 99DE-01040832.
PR

XX 31-AUG-1999; 99DE-01041378.
PR

XX 31-AUG-1999; 99DE-01041379.
PR

XX 31-AUG-1999; 99DE-01041380.
PR

XX 31-AUG-1999; 99DE-01041394.
PR

XX 03-SEP-1999; 99DE-01041396.
PR

XX 03-SEP-1999; 99DE-01042076.
PR

XX 03-SEP-1999; 99DE-01042077.
PR

XX 03-SEP-1999; 99DE-01042079.
PR

XX 03-SEP-1999; 99DE-01042086.
PR

XX 03-SEP-1999; 99DE-01042087.
PR

XX 03-SEP-1999; 99DE-01042088.
PR

XX 03-SEP-1999; 99DE-01042095.
PR

XX 03-SEP-1999; 99DE-01042124.
PR

XX 03-SEP-1999; 99DE-01042129.
PR

XX 09-MAR-2000; 2000US-0189770P.
PR

XX (SADI) BASF AG.
PA

XX Pompejus M, Kroeger B, Schroeder H, Zeider O, Haberhauer G;
PI

XX WPI; 2001-137957/14.
XX

XX N-PSDB; AAF71872.
DR

XX Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway proteins, useful for producing fine chemicals in microorganisms, including organic acids, nonproteinogenic amino acids, and purine and pyrimidine bases.
PT

XX Claim 20; Page 505-507; 1737pp; English.
PS
XX
CC AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
CC nucleic acids are useful for the production of fine chemicals in
CC microorganisms, including organic acids, nonproteinogenic amino acids,
CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,
CC vitamins, cofactors, polyketides and enzymes
XX
SQ Sequence 568 AA;
Query Match 9.5%; Score 7; DB 4; Length 568;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 37 ALTAVALS 43
DB 10 ALTAVALS 16
RESULT 11
AAB80200 ID AAB80200 standard; protein; 568 AA.
XX AC AAB80200;
XX
DT 30-APR-2001 (first entry)
XX
DE Corynebacterium glutamicum MP protein sequence SEQ ID NO:1134.
XX
KW Corynebacterium glutamicum; metabolic pathway protein; MP protein;
KW fine chemical production; microorganism; organic acid; nucleoside;
KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleotide;
KW lipid; saturated fatty acid; unsaturated fatty acid; diol; vitamin;
KW carbohydrate; aromatic compound; cofactor; polyketide; enzyme.
XX
OS Corynebacterium glutamicum.
FN WO200100843-A2.
XX
XX 04-JAN-2001.
XX
XX 23-JUN-2000; 2000WO-IB000923.
XX
XX 25-JUN-1999; 99US-0141031P.
PR 01-JUL-1999; 99DE-01030476.
PR 02-JUL-1999; 99US-0142101P.
PR 08-JUL-1999; 99DE-01031415.
PR 08-JUL-1999; 99DE-01031418.
PR 08-JUL-1999; 99DE-01031419.
PR 08-JUL-1999; 99DE-01031420.
PR 08-JUL-1999; 99DE-01031424.
PR 08-JUL-1999; 99DE-01031428.
PR 08-JUL-1999; 99DE-01031434.
PR 08-JUL-1999; 99DE-01031435.
PR 08-JUL-1999; 99DE-01031443.
PR 08-JUL-1999; 99DE-01031453.
PR 08-JUL-1999; 99DE-01031457.
PR 08-JUL-1999; 99DE-01031465.
PR 08-JUL-1999; 99DE-01031478.
PR 08-JUL-1999; 99DE-01031510.
PR 08-JUL-1999; 99DE-01031541.
PR 08-JUL-1999; 99DE-01031543.
PR 08-JUL-1999; 99DE-01031573.
PR 08-JUL-1999; 99DE-01031632.
PR 08-JUL-1999; 99DE-01031634.
PR 08-JUL-1999; 99DE-01031636.
PR 09-JUL-1999; 99DE-01032125.
PR 09-JUL-1999; 99DE-01032126.
PR 09-JUL-1999; 99DE-01032130.
PR 09-JUL-1999; 99DE-01032186.

PR 09-JUL-1999; 99DE-01032206.
PR 09-JUL-1999; 99DE-01032227.
PR 09-JUL-1999; 99DE-01032228.
PR 09-JUL-1999; 99DE-01032229.
PR 09-JUL-1999; 99DE-01032230.
PR 14-JUL-1999; 99DE-01032922.
PR 14-JUL-1999; 99DE-01032926.
PR 14-JUL-1999; 99DE-01032928.
PR 14-JUL-1999; 99DE-01033004.
PR 14-JUL-1999; 99DE-01033005.
PR 14-JUL-1999; 99DE-01033006.
PR 12-AUG-1999; 99US-0148613P.
PR 27-AUG-1999; 99DE-01040764.
PR 27-AUG-1999; 99DE-01040765.
PR 27-AUG-1999; 99DE-01040766.
PR 27-AUG-1999; 99DE-01040832.
PR 31-AUG-1999; 99DE-01041378.
PR 31-AUG-1999; 99DE-01041379.
PR 31-AUG-1999; 99DE-01041380.
PR 31-AUG-1999; 99DE-01041394.
PR 31-AUG-1999; 99DE-01041396.
PR 03-SEP-1999; 99DE-01042077.
PR 03-SEP-1999; 99DE-01042079.
PR 03-SEP-1999; 99DE-01042086.
PR 03-SEP-1999; 99DE-01042087.
PR 03-SEP-1999; 99DE-01042088.
PR 03-SEP-1999; 99DE-01042095.
PR 03-SEP-1999; 99DE-01042124.
PR 03-SEP-1999; 99DE-01042129.
PR 09-MAR-2000; 2000US-0187970P.
XX
XX (BADI) BASF AG.
PA
XX Pompejus M. Kroeger B, Schroeder H, Zelder O, Haberhauer G;
PI WPI: 2001-137957/14.
XX N-PSDB; AAF72319.
DR
XX
XX Nucleic acids from Corynebacterium glutamicum encoding metabolic pathway
PT proteins, useful for producing fine chemicals in microorganisms,
PT including organic acids, nonproteinogenic amino acids, and purine and
PT pyrimidine bases.
XX
PS Claim 20; Page 1698-1700; 1737pp; English.
XX
XX AAF71753 to AAF72330 encode the Corynebacterium glutamicum metabolic
CC pathway (MP) proteins given in AAB79634 to AAB80211. The C. glutamicum MP
CC nucleic acids are useful for the production of fine chemicals in
CC microorganisms, including organic acids, nonproteinogenic amino acids,
CC purine and pyrimidine bases, nucleosides, nucleotides, lipids, saturated
CC and unsaturated fatty acids, diols, carbohydrates, aromatic compounds,
CC vitamins, cofactors, polyketides and enzymes
XX
SQ Sequence 568 AA;
Query Match 9.5%; Score 7; DB 4; Length 568;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 37 ALTAVALS 43
DB 10 ALTAVALS 16
RESULT 12
AAU71921 ID AAU71921 standard; protein; 568 AA.
XX
XX AAU71921;
XX
XX 26-FEB-2002 (first entry)
XX

C. glutamicum metabolic pathway protein encoded by gene #56.
Metabolic pathway protein; MP; lysine biosynthesis pathway;
methionine biosynthesis pathway; large-scale production of fine chemical;
Corynebacterium diphtheriae; diphtheria.
Corynebacterium glutamicum.
WO200166573-A2.
13-SEP-2001.
22-DEC-2000; 2000WO-IB002035.
09-MAR-2000; 2000US-0187970P.
23-JUN-2000; 2000US-00606740.
(BADI) BASF AG.
Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
Kim J, Lee H, Hwang B;
WPI; 2001-582269/65.
N-PSDB; AAS96131.
Nucleic acids encoding metabolic pathway proteins from Corynebacterium
glutamicum, useful for producing methionine and lysine in Corynebacterium
and Brevibacterium.
Disclosure; Page 307-309; 316pp; English.
The present invention relates to the isolation of novel Corynebacterium
glutamicum genes (AAS96073-AAS96132) encoding metabolic pathway (MP)
proteins. The metabolic pathway proteins of the invention include enzymes
involved in the lysine and methionine biosynthetic pathways. The
polynucleotide sequences of the invention can be used for the large-scale
production and/or modulation of expression of fine chemicals such as
lysine and methionine. The sequences of the invention may be used to
identify C. glutamicum and related organisms e.g. C. diphtheriae in a
subject to detect diphtheria. AAU71863-AAU71922 represent the novel C.
glutamicum metabolic pathway proteins of the invention
Sequence 568 AA;
Query Match 9.5%; Score 7; DB 4; Length 568;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 37 ALTAVALS 43
Db 10 ALTAVALS 16
RESULT 13
AAG90589
ID AAG90589 standard; protein; 568 AA.
AC AAG90589;
XX 26-SEP-2001 (first entry);
XX C glutamicum protein fragment SEQ ID NO: 4343.
XX Corynebacterium; amino acid synthesis; vitamin; saccharide;
KW organic acid synthesis.
XX Corynebacterium glutamicum.
OS EP1108790-A2.
PN 20-JUN-2001.
XX 18-DEC-2000; 2000EP-00127688.
C. glutamicum metabolic pathway protein encoded by gene #56.
Metabolic pathway protein; MP; lysine biosynthesis pathway;
methionine biosynthesis pathway; large-scale production of fine chemical;
Corynebacterium diphtheriae; diphtheria.
Corynebacterium glutamicum.
WO200166573-A2.
13-SEP-2001.
22-DEC-2000; 2000WO-IB002035.
09-MAR-2000; 2000US-0187970P.
23-JUN-2000; 2000US-00606740.
(BADI) BASF AG.
Pompejus M, Kroeger B, Schroeder H, Zelder O, Haberhauer G;
Kim J, Lee H, Hwang B;
WPI; 2001-582269/65.
N-PSDB; AAS96131.
Nucleic acids encoding metabolic pathway proteins from Corynebacterium
glutamicum, useful for producing methionine and lysine in Corynebacterium
and Brevibacterium.
Disclosure; Page 307-309; 316pp; English.
The present invention relates to the isolation of novel Corynebacterium
glutamicum genes (AAS96073-AAS96132) encoding metabolic pathway (MP)
proteins. The metabolic pathway proteins of the invention include enzymes
involved in the lysine and methionine biosynthetic pathways. The
polynucleotide sequences of the invention can be used for the large-scale
production and/or modulation of expression of fine chemicals such as
lysine and methionine. The sequences of the invention may be used to
identify C. glutamicum and related organisms e.g. C. diphtheriae in a
subject to detect diphtheria. AAU71863-AAU71922 represent the novel C.
glutamicum metabolic pathway proteins of the invention
Sequence 568 AA;
Query Match 9.5%; Score 7; DB 4; Length 568;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 37 ALTAVALS 43
Db 10 ALTAVALS 16
RESULT 14
ABB59908
ID ABB59908 standard; protein; 647 AA.
XX ABB59908;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster polypeptide SEQ ID NO 6516.
XX Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX Drosophila melanogaster.
OS WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li FWD, Myers EW;
XX WPI; 2001-656860/75.
DR N-PSDB; ABL04011.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2004, 21:57:35 ; Search time 325.49 Seconds
(without alignments)
2443.166 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 1095

Sequence: 1 MCNTNMSVPTGAVTTSQIP.....ALCVIREICCRSSSSSTG 216

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2432557 seqs, 1840798884 residues

Total number of hits satisfying chosen parameters: 4865114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10057510/runat_09032004.162236.26038/app_query.fasta_1.654
-DS=Published Applications NA -QWTF=fastap -SUFFIX=std.rnpb -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNIQ=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500
-MAXLEN=2000000000 -USER=US10057510.cgn 1.1 213 @runat_09032004.162236.26038
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PTCUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

1	1084.5	99.0	652	13	US-10-057-510-3	Sequence 3, Appli
2	1084.5	99.0	1476	10	US-09-029-327-1	Sequence 1, Appli
3	1084.5	99.0	2372	9	US-09-752-983-1	Sequence 1, Appli
4	1084.5	99.0	2372	9	US-09-956-425-7	Sequence 7, Appli
5	1084.5	99.0	2372	9	US-09-851-771A-1	Sequence 1, Appli
6	1084.5	99.0	2372	10	US-09-541-848-1	Sequence 1, Appli
7	1084.5	99.0	2372	10	US-09-873-367C-450	Sequence 450, App
8	1084.5	99.0	2372	12	US-10-232-951-34	Sequence 34, Appl
9	1084.5	99.0	2372	14	US-10-007-926A-121	Sequence 121, App
10	1084.5	99.0	2372	15	US-10-005-344-1	Sequence 1, Appli
11	1084.5	99.0	2372	15	US-10-422-536-136	Sequence 136, App
12	805	73.5	1470	15	US-10-005-344-324	Sequence 324, App
13	804.5	73.5	1710	9	US-09-956-425-5	Sequence 5, Appli
14	804.5	73.5	1710	10	US-09-541-848-12	Sequence 12, Appl
15	290.5	26.5	176	14	US-10-211-088-144	Sequence 144, App
16	236.5	21.6	381	10	US-09-918-995-36647	Sequence 36647, A
17	214	19.5	199	13	US-10-109-213-1	Sequence 1, Appli
18	214	19.5	199	13	US-10-109-213-2	Sequence 2, Appli
19	127	11.6	755	12	US-10-424-599-126709	Sequence 126709, S
20	125	11.4	73	10	US-09-541-848-49	Sequence 49, Appl
21	119	10.9	500	9	US-09-752-983-2	Sequence 2, Appli
22	119	10.9	500	9	US-09-851-771A-2	Sequence 2, Appli
23	119	10.9	500	15	US-10-005-344-2	Sequence 5, Appli
24	100	9.1	10640	14	US-10-304-095-5	Sequence 5, Appli
25	99.5	8.9	1302	8	US-08-945-038-5	Sequence 5, Appli
26	97.5	8.9	2058	12	US-10-424-599-119795	Sequence 119795, S
27	93.5	8.5	952	12	US-10-425-114-14558	Sequence 14558, A
28	92.5	8.4	4638	10	US-09-814-353-21179	Sequence 21179, A
29	91	8.3	1539	12	US-10-282-122A-22970	Sequence 22970, A
30	91	8.3	2012	12	US-10-425-114-6428	Sequence 6428, AP
31	91	8.3	5547	9	US-09-880-107-1715	Sequence 1715, AP
32	90.5	8.3	1385	12	US-10-425-114-26817	Sequence 26817, A
33	90	8.2	2010	15	US-10-369-493-34074	Sequence 34074, A
34	89.5	8.2	905	12	US-10-424-599-135024	Sequence 135024, S
35	89.5	8.2	2644	14	US-10-128-714-1444	Sequence 1444, AP
36	89.5	8.2	2644	14	US-10-128-714-6444	Sequence 6444, AP
37	89.5	8.2	4644	14	US-10-128-714-444	Sequence 444, App
38	89.5	8.2	4644	14	US-10-128-714-5444	Sequence 5444, AP
39	89	8.1	2812	14	US-10-198-846-9693	Sequence 9693, App
40	89	8.1	3027	14	US-10-078-531-1	Sequence 1, Appli
41	88.5	8.1	3648	9	US-09-775-181-1	Sequence 1, Appli
42	88.5	8.1	3687	14	US-10-232-539-1	Sequence 5, Appli
43	88.5	8.1	7148	9	US-09-775-181-5	Sequence 1, Appli
44	88	8.0	1314	14	US-10-176-584A-1	Sequence 534, AP
45	88	8.0	2215	12	US-10-425-114-6334	Sequence 6334, AP

ALIGNMENTS

RESULT 1

US-10-057-510-3
; Sequence 3, Application US/10057510
; Publication No. US2002009580A1
; GENERAL INFORMATION:
; APPLICANT: Mandabalan, Krishnan
; APPLICANT: Yang, Meljia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDM INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT APPLICATION NUMBER: US/10/057,510
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: USSN 09/510,252
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USSN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USSN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 652
; TYPE: DNA

```
ORGANISM: Homo sapiens
US-10-057-510-3
Alignment Scores:
Pred. No.: 8,78e-131 Length: 652
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 13 Gaps: 1

US-10-057-510-4 (1-216) x US-10-057-510-3 (1-652)
Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 1 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTAAACCACTCACAGATTCCA 60
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLeu 40
Db 61 GCTTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120
Qy 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyLeu---GlnTyr 59
Db 121 GTTGTGTCACAAAAGACACTTATCTATGAAAGAGGTTCTTTTATCTTGCCAGTAT 180
Qy 60 IleMetThrLysArgLeuTyAspGluGlnGlnHisIleValTyrCysSerAsnAsp 79
Db 181 ATTATGACTAAACGATTATATGATGAGAAGCAACATATTGTATATTGTTCAATGAT 240
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 241 CTTCAGGAGATTCTTTGGCGTGCCAAAGCTTCTCTGTGAAGAGACACAGGAAATATAT 300
Qy 100 ThrMetIleTyrArgAsnLeuValValValAsnGlnGlnGlnGlnGlnGlnGln 119
Db 301 ACCATGATCTACAGGAACCTTGTTAGTAGTATCATGACGAGGAATCATCGGACTCAGGTACA 360

US-09-029-327-1
Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 1 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTAAACCACTCACAGATTCCA 60
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLeu 40
Db 61 GCTTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120
Qy 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyLeu---GlnTyr 59
Db 121 GTTGTGTCACAAAAGACACTTATCTATGAAAGAGGTTCTTTTATCTTGCCAGTAT 180
Qy 60 IleMetThrLysArgLeuTyAspGluGlnGlnHisIleValTyrCysSerAsnAsp 79
Db 181 ATTATGACTAAACGATTATATGATGAGAAGCAACATATTGTATATTGTTCAATGAT 240
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 241 CTTCAGGAGATTCTTTGGCGTGCCAAAGCTTCTCTGTGAAGAGACACAGGAAATATAT 300
Qy 100 ThrMetIleTyrArgAsnLeuValValValAsnGlnGlnGlnGlnGlnGlnGln 119
Db 301 ACCATGATCTACAGGAACCTTGTTAGTAGTATCATGACGAGGAATCATCGGACTCAGGTACA 360

US-10-057-510-4 (1-216) x US-09-029-327-1 (1-1476)
Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 1 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTAAACCACTCACAGATTCCA 60
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLeu 40
Db 61 GCTTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120
Qy 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyLeu---GlnTyr 59
Db 121 GTTGTGTCACAAAAGACACTTATCTATGAAAGAGGTTCTTTTATCTTGCCAGTAT 180
Qy 60 IleMetThrLysArgLeuTyAspGluGlnGlnHisIleValTyrCysSerAsnAsp 79
Db 181 ATTATGACTAAACGATTATATGATGAGAAGCAACATATTGTATATTGTTCAATGAT 240
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 241 CTTCAGGAGATTCTTTGGCGTGCCAAAGCTTCTCTGTGAAGAGACACAGGAAATATAT 300
Qy 100 ThrMetIleTyrArgAsnLeuValValValAsnGlnGlnGlnGlnGlnGlnGln 119
Db 301 ACCATGATCTACAGGAACCTTGTTAGTAGTATCATGACGAGGAATCATCGGACTCAGGTACA 360

US-09-029-327-1
Alignment Scores:
Pred. No.: 3,22e-130 Length: 1476
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 13 Gaps: 1

US-10-057-510-4 (1-216) x US-09-029-327-1 (1-1476)
Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 1 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTAAACCACTCACAGATTCCA 60
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLeu 40
Db 61 GCTTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120
Qy 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyLeu---GlnTyr 59
Db 121 GTTGTGTCACAAAAGACACTTATCTATGAAAGAGGTTCTTTTATCTTGCCAGTAT 180
Qy 60 IleMetThrLysArgLeuTyAspGluGlnGlnHisIleValTyrCysSerAsnAsp 79
Db 181 ATTATGACTAAACGATTATATGATGAGAAGCAACATATTGTATATTGTTCAATGAT 240
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 241 CTTCAGGAGATTCTTTGGCGTGCCAAAGCTTCTCTGTGAAGAGACACAGGAAATATAT 300
Qy 100 ThrMetIleTyrArgAsnLeuValValValAsnGlnGlnGlnGlnGlnGlnGln 119
Db 301 ACCATGATCTACAGGAACCTTGTTAGTAGTATCATGACGAGGAATCATCGGACTCAGGTACA 360

US-09-029-327-1
Sequence 1, Application US/09029327
Publication No. US20030060432A1
GENERAL INFORMATION:
APPLICANT: TOCQUE, Bruno
APPLICANT: WASLYK, Bohdan
APPLICANT: DUBS-POTERSZMAN,
APPLICANT: Marie-Christine
TITLE OF INVENTION: ANTAGONISTS OF THE ONCOGENIC ACTIVITY OF
TITLE OF INVENTION: THE PROTEIN MDM2, AND USE THEREOF IN THE TREATMENT OF
TITLE OF INVENTION: CANCERS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Road, Mailstop 3c43
```

QY 120 SerValSerGluAenArgCysHisLeuGluGlyVlySerAspGlnLysAspLeuValGln 139
DB 361 TCTGTGAGTGAGACAGGTGTACCTTGAAGTGGAGTGATCAAAAGGACCTTGTACAA 420
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
DB 421 GAGCTTCAGGAAGAGAACTTCATCTTACATTTGGTTCTAGACCATCTACCTCATCT 480
QY 160 ArgAcArgAlaIleSerGluThrGluAenSerAspGluLeuSerGlyGluArgGln 179
DB 481 AGAAGGAGAGCAATAGTGAGACAGAGAAATTCAGATGAATTCCTGGTGAAGACAA 540
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 541 AGAAACGCCACAACTCATAGTATTTCCCTTCTCTTGTGATGAAGCCCTGGCTCTGTGT 600
QY 200 ValIleArgGluLeuCysCysGluArgSerSerSerSerSerGluSerThrGly 216
DB 601 GTAATAAGGAGATATGTTGTGAAGAGACAGTAGCAGTGAATCTACAGGG 651

RESULT 3

US-09-752-983-1
; Sequence 1, Application US/09752983
; Patent No. US20010016575A1
; GENERAL INFORMATION:
; APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.
; APPLICANT: Graham, Brett P. Monia
; TITLE OF INVENTION: ANTISENSE MODULATION OF HUMAN MDM2
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 271
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Jane Massey Licata
; STREET: 66 East Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: U.S.A.
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
; OPERATING SYSTEM: WINDOWS 95
; SOFTWARE: WORDPERFECT 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/752,983
; FILING DATE: 02-Jan-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/280,805
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Licata, Jane Massey
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0346
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 609-810-1515
; TELEFAX: 609-810-1454
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2372 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Unknown
; ANTI-SENSE: NO
; PUBLICATION INFORMATION:
; AUTHORS: Oliner, J.D.
; AUTHORS: Kinzler, K.W.
; AUTHORS: Meltzer, P.S.
; AUTHORS: George, D.L.
; AUTHORS: Vogelstein, B.
; TITLE: Amplification of a gene encoding a
; TITLE: p53-associated protein in human sarcomas
; JOURNAL: Nature
; VOLUME: 358

ISSUE: 6381
; PAGES: 80-83
; DATE: 02-JUL-1992
US-09-752-983-1
Alignment Scores:
Pred. No.: 6.84e-130 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 9 Gaps: 1
US-10-057-510-4 (1-216) x US-09-752-983-1 (1-2372)
QY 1 MetCysAenThrAenMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACCTACTGTAGTGGTGTGTACCACTCACGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysLeuLysSer 40
DB 372 GCTTCGGACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 432 GTTGGTGCACAAAAGACACTTATACCTATGAAAGAGGCTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 492 ATTATGACTTAACGATTATATGATGAGAGACACACATATTGTATATTGTTCAAAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTCTAGGAGATTTGTTGGGTGCCAAGCTTCTCTGTGAAGAGACAGCAAGAAATATAT 611
QY 100 ThrMetIleTyrArgAenLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
DB 612 ACCATGATCTACAGGAACCTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAenArgCysHisLeuGluGlyVlySerAspGlnLysAspLeuValGln 139
DB 672 TCTGTGAGTGAGACAGGTGTACCTTGAAGTGGAGTGATCAAAAGGACCTTGTACAA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
DB 732 GAGCTTCAGGAAGAGAAACCTTCATCTTCATTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgAlaIleSerGluThrGluAenSerAspGluLeuSerGlyGluArgGln 179
DB 792 AGAAGGAGAGCAATAGTGAGACAGAGAAATTCAGATGAATTCATCTGGTGAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 852 AGAAACGCCACAAATCTGATAGTATTTCCCTTCTCTTGTATGAAGCCCTGGCTCTGTGT 911
QY 200 ValIleArgGluLysCysGluArgSerSerSerSerSerGluSerThrGly 216
DB 912 GTAATAAGGAGATATGTTGTGAAGAGACAGTAGCAGTGAATCTACAGGG 962
RESULT 4
US-09-956-425-7
; Sequence 7, Application US/09956425
; Patent No. US20020045192A1
; GENERAL INFORMATION:
; APPLICANT: Kriwacki, Richard
; APPLICANT: Bothner, Brian
; APPLICANT: Lewis, William
; TITLE OF INVENTION: Arf and Hdm2 Interaction Domains and Method of Use Thereof
; FILE REFERENCE: 1340/1/035
; CURRENT APPLICATION NUMBER: US/09/956,425
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.11

SEQ ID NO 7
LENGTH: 2372
TYPE: DNA
ORGANISM: Homo sapiens
US-09-956-425-7

Alignment Scores:

Pred. No.: 6,84e-130 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 9 Gaps: 1

US-10-057-510-4 (1-216) x US-09-956-425-7 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACCTACTGATGGTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysLeuLeuLysSer 40
DB 372 GCTTCGGAACAGAGACCTGGTGTAGACCAAGCCATGCTTTTGAAGTTATTAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 432 GTTGGTGCAACAAAAGACACTTACTATGAAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisLeuValTyrCysSerAsnAsp 79
DB 492 ATTATGCTAAACGATTATGATGTAGAGCAACACATATTGATTTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTTCAGGAGATTGTTTGGCGTGCCAAAGCTTCTCTGTGAAAGACACAGGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnLysSerSerAspSerGlyThr 119
DB 612 ACCATGATCTACAGGAATCTGGTAGTAGTCAATCAGCAGGAATCTCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGluGlySerAspGlnLysAspLeuValGln 139
DB 672 TCTGTGATGAGAACAGGTGTCCTTGAAGGTGGGAGTGATCAAAAGGACCTTGACAA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
DB 732 GAGCTTCAGGAAGAGAAACCTTCATCTTCACATTGCTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
DB 792 AGAAGGAGAGCAATTAGTGACAGACAGAGAAATTCAGATGAATATCTGGTGACACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 852 AGAAACGCCACAAATCTGATAGTATTTCCCTTTCCCTTTCATGAAAGCCCTGCTGTGT 911
QY 200 ValIleArgGluLeuLysCysGluArgSerSerSerSerGluSerThrGly 216
DB 912 GTAATAAGGGAGATATGTTGTGAAGAAGCAGTAGCAGTGAATCTACAGGG 962

RESULT 5

US-09-851-771A-1
Sequence 1, Application US/09851771A
Patent No. US200201511A1
GENERAL INFORMATION:
APPLICANT: Loren J. Miraglia, Pamela Nero, Mark J.
Graham, Brett P. Monia
TITLE OF INVENTION: ANTIGENSE OLIGONUCLEOTIDE
MODULATION OF HUMAN MDM2 EXPRESSION
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Jane Massey Licata
STREET: 66 East Main Street

CITY: Marlton
STATE: NJ
COUNTRY: U.S.A.
ZIP: 08053
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/851,771A
FILING DATE: 09-May-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/048,810
FILING DATE: 1998-03-26
ATTORNEY/AGENT INFORMATION:
NAME: Licata, Jane Massey
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: ISPH-0302
TELECOMMUNICATION INFORMATION:
TELEPHONE: 609-779-2400
TELEFAX: 609-810-1454
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2372 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Unknown
ANTI-SENSE: No
PUBLICATION INFORMATION:
AUTHORS: Oliner,J.D.,Kinzler,K.W.,Meltzer,P.S.,George,D.L.,Vogelstein,B.
TITLE: Amplification of a gene encoding a p53-associated protein in hu

US-09-851-771A-1

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Alignment Scores:
Pred. No.: 6,84e-130 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 9 Gaps: 1

US-10-057-510-4 (1-216) x US-09-851-771A-1 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACCTACTGATGGTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysSer 40
DB 372 GCTTCGGAACAGAGACCTGGTGTAGACCAAGCCATGCTTTTGAAGTTATTAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 432 GTTGGTGCAACAAAAGACACTTACTATGAAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisLeuValTyrCysSerAsnAsp 79
DB 492 ATTATGCTAAACGATTATGATGTAGAGCAACACATATTGATTTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTTCAGGAGATTGTTTGGCGTGCCAAAGCTTCTCTGTGAAAGACACAGGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnLysSerSerAspSerGlyThr 119

Db	612	ACCATGATCTACAGAACTTGGTAGTAGTCAATCAGCAGGAATCATCGACTCAGGTACA	671
QY	120	SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnIysAspLeuValGln	139
Db	672	TCTGTGACGTAGAAACAGGTGTCACTTGAAGGTGGAGTGATCAAAAGGACCTTGTACAA	731
QY	140	GluLeuGlnGluIysProSerSerHisLeuValSerArgProSerThrSerSer	159
Db	732	GAGCTTCAGGAACAGAACTTCATCTTTCACATTTGGTTTCTAGACCATCTACTCATCT	791
QY	160	ArgArgArgAlaIleSerGluThrGluIuAsnSerAspGluLeuSerGlyGluArgGln	179
Db	792	AGAAAGCAGACGAATTAGTGCAGACAGAAATAATTCAGATGAATATCTGCTGAAACGACAA	851
QY	180	ArgLysArgHisIleValSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys	199
Db	852	AGAAACAGCCCAAAATCTGATAGTATTTCCTCTTTCCTTTCATGAAAGCTGGCTCTGTCT	911
QY	200	ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly	216
Db	912	GTAATAAGGAGATATTGTTGCAAGAACAGTAGTACGAGTGAATCTACAGGG	962

RESULT 6

```

US-09-541-848--1
; Sequence 1, Application US/09541848
; Publication No. US20030119765A1
; GENERAL INFORMATION:
; APPLICANT: CHEN, Jtandong
; APPLICANT: AGRAWAL, Sudhir
; APPLICANT: ZHANG, Ruiwen
; TITLE OF INVENTION: MD-M2-SPECIFIC ANTISENSE OLIGONUCLEOTIDES
; FILE REFERENCE: 29924/98057C
; CURRENT APPLICATION NUMBER: US/09/541,848
; CURRENT FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: 09/383,507
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 09/073,567
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 08/916,834
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-541-848--1

```

```

Alignment Scores:
Pred. NO.:      6.84e-130
Score:          1081.50
Percent Similarity: 99.54%
Best Local Similarity: 99.54%
Query Match:    99.04%
Gaps:           10
DB:             1
Length:         2372
Matches:        216
Conservative:   0
Mismatch:       0
Indels:         1
Gaps:           1

```

US-10-057-510-4 (1-216) X US-09-541-848-1 (1-2372)

QY		1	MetCysAsnThrAsnMetSerValProThraspGlyAlaValThrThrSerGlnIlePro	20
Db		312	ATGTGCAATTACCAACATCTCTGTACTTACTGATGGTGCTGTAAACCACCTCACAGATTCCA	371
QY		21	AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLysSer	40
Db		372	GCTTCGGAACCAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTGAAGTTATTATAAAGTCT	431
QY		41	ValGlyAlaGlnLysAspThr.TyThr.MetLysGluValLeuPheTyrlxLeu---GlnTyr	59
Db		432	GTGTGTGCACAAAAGACACTTATACATGTAAGAGGCTTCTTTTTTATCTTGCCCAAGAT	491
QY		60	IleMetThrLysArgLeuTyAspGluLysGlnHisIleValTyrCysSerAsnAsp	79
Db		492	ATTATGACTTAAACCATATATGATGAGAAGCAACAACATATTGTATATTGTTCAATATGAT	551

QY	80	LeuLeuGlyAspLeuPheClyValProSerPheSerValLysGluHisArgLysIleThr	99
DB	552	CTTCTAGGAGATTGTGTTGGCGTGCACAGCTTCTCTGTGAAAGAGACACAGAGAAATATAT	611
QY	100	ThrMetIleTyArgAsnLeuValValAenGlnGlnGluSerSerAspSerGlyThr	119
DB	612	ACCATGATCTACAGAACCTTGGTAGTAGTCTAAATCAGCAGGAATCATCGGATCTCAGGTACA	671
QY	120	SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln	139
DB	672	TCTGTGAGTGAGAACAGGCTGTCACTTGAAGGTGGAGTGATCAAAAGAGACCTTGTACAA	731
QY	140	GluLeuGlnGluCluLysProSerSerSerHisIleuValSerArgProSerThrSerSer	159
DB	732	GAGCTTCAGGAAGAGAAACCTTCATCTTCATTTGGTTCTTAGACCATCTACCTCATCT	791
QY	160	ArgArgArgAlaIleSerGluThrGluGluAenSerAspGluLeuSerGlyGluArgGln	179
DB	792	AGAAGGAGGCAATTAGTGAGNACAGAGAAATTCAGATGAATTATCTGGTGAACGACAA	851
QY	180	ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys	199
DB	852	AGAAACGCCCAAAATCTGATAGTATTTCCCTTTTCCTTTGATGAAGCGCTGGCTCTGTGT	911
QY	200	ValIleArgGluIleCysCysGluArgSerSerSerGluSerThrGly	216
DB	912	GTAATAAGGGAGATATGTTGTGAAAGACAGCATGACAGTGAATCTACAGG	962

RESULT 7

```

US-09-873-367C-450
/ Sequence 450, Application US/09873367C
/ Publication No. US20030165839A1
/ GENERAL INFORMATION:
/ APPLICANT: Young, Paul
/ APPLICANT: Soppet, Daniel
/ APPLICANT: Andrews, Gregory
/ APPLICANT: Augustus, Meena
/ APPLICANT: Ebner, Reinhard
/ APPLICANT: Carter, Kenneth
/ TITLE OF INVENTION: Cancer Gene Determinant
/ TITLE OF INVENTION: Signature Gene Sets
/ FILE REFERENCE: 689290-64
/ CURRENT APPLICATION NUMBER: US/09/873,367C
/ CURRENT FILING DATE: 2003-04-29
/ PRIOR APPLICATION NUMBER: U.S. 60/236,891
/ PRIOR FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: U.S. 60/236,842
/ PRIOR FILING DATE: 2000-09-29
/ PRIOR APPLICATION NUMBER: U.S. 60/244,867
/ PRIOR FILING DATE: 2000-11-01
/ PRIOR APPLICATION NUMBER: U.S. 60/245,084
/ PRIOR FILING DATE: 2000-11-01
/ NUMBER OF SEQ ID NOS: 1067
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 450
/ LENGTH: 2372
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-873-367C-450

```

Alignment Scores:	
Pred. No.:	6.84e-130
Score:	1084.50
Length:	2372
Matches:	216
Conservative:	0
Mismatches:	0
Indels:	1
Gaps:	1
DB:	10
Query Match:	99.04%
Best Local Similarity:	99.54%
Percent Similarity:	99.54%
Score:	1084.50

US-10-057-510-4 (1-216) X US-09-873-367C-450 (1-2372)

Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20


```
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: mouse double minute 2, human homolog of;
; OTHER INFORMATION: p53-binding protein (MDM2) gene.
US-10-007-926A-121

Alignment Scores:
Pred. No.: 6,84e-130 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 14 Gaps: 1

US-10-057-510-4 (1-216) x US-10-007-926A-121 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCANTACCAACATGCTGTACCTACTGATGGTGTGTAAACCACTCCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysLeuLysSer 40
Db 372 GCTTCGGACACAGAGACCCCTGGTAGACCAAGCCATTGCTTTTGAAGTTATTAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 432 GTTGGTGCCACAAAAGACACTTACTATGAAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisLeuValTyrCysSerAsnAsp 79
Db 492 ATTATGACTAAACGATTATATGATGAGACCAACAACTATTGTTATTTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 552 CTTCCTAGGAGATTGTTGGCGTCCCAAGCTTCTCTGTGAAAGAGACACAGGAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGGAACCTTGGTAGTCACTATGAGGAGGAGTATCAAAAGAGACCTTGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTGAGTGAGAACACAGTGTACCTTGAAGGTGGAGTATCAAAAGAGACCTTGTACA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GAGCTTCAGGAGAGAAACCTTCATCTTCACATTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluLysAsnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATATTAGTGAGACAGAGAAATTCAGATGAATATCTGTGTGAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAAACGCCACAAATCTGATAGTATTTCCCTTTCCCTTTGATGAAGCCTTGGCTCTGTGT 911
QY 200 ValIleArgGluIleCysGluArgSerSerSerSerSerGluSerThrGly 216
Db 912 GTAATAAGGGAGATATGTTGTGAAAGAGACAGTACAGTGAATCTACAGGG 962

RESULT 10
US-10-005-344-1
; Sequence 1, Application US/10005344
; Publication No. US2003020362A1
; GENERAL INFORMATION:
; APPLICANT: Loren J. Miraglia
; APPLICANT: Pamela Nero
; APPLICANT: Mark J. Graham
; APPLICANT: Brett P. Monia
; APPLICANT: Erich Koller
; APPLICANT: Mingyi Chiang
```

```
; APPLICANT: Mano Manoharan
; TITLE OF INVENTION: Antisense Modulation of mdm2 expression.
; FILE REFERENCE: ISPH-0622
; CURRENT APPLICATION NUMBER: US/10/005,344
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 09/048,810
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: US 09/280,805
; PRIOR FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 379
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (312)...(1787)
US-10-005-344-1

Alignment Scores:
Pred. No.: 6,84e-130 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 15 Gaps: 1

US-10-057-510-4 (1-216) x US-10-005-344-1 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCANTACCAACATGCTGTACCTACTGATGGTGTGTAAACCACTCCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysLeuLysSer 40
Db 372 GCTTCGGACACAGAGACCCCTGGTAGACCAAGCCATTGCTTTTGAAGTTATTAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 432 GTTGGTGCCACAAAAGACACTTACTATGAAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisLeuValTyrCysSerAsnAsp 79
Db 492 ATTATGACTAAACGATTATATGATGAGACCAACAACTATTGTTATTTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 552 CTTCCTAGGAGATTGTTGGCGTCCCAAGCTTCTCTGTGAAAGAGACACAGGAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGGAACCTTGGTAGTCACTATGAGGAGGAGTATCAAAAGAGACCTTGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTGAGTGAGAACACAGTGTACCTTGAAGGTGGAGTATCAAAAGAGACCTTGTACA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GAGCTTCAGGAGAGAAACCTTCATCTTCACATTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluLysAsnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATATTAGTGAGACAGAGAAATTCAGATGAATATCTGTGTGAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAAACGCCACAAATCTGATAGTATTTCCCTTTCCCTTTGATGAAGCCTTGGCTCTGTGT 911
QY 200 ValIleArgGluIleCysGluArgSerSerSerSerSerGluSerThrGly 216
Db 912 GTAATAAGGGAGATATGTTGTGAAAGAGACAGTACAGTGAATCTACAGGG 962
```

```
RESULT 11
US-10-422-536-136
; Sequence 136, Application US/10422536
; Publication No. US20040014100A1
; GENERAL INFORMATION:
; APPLICANT: Kinella, Todd
; APPLICANT: Lorens, James
; APPLICANT: Pray, Todd
; APPLICANT: Bennett, Mark
; TITLE OF INVENTION: IN VIVO PRODUCTION OF CYCLIC PEPTIDES FOR INHIBITING
; TITLE OF INVENTION: PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: A-71433-1/AMF/CVO
; CURRENT APPLICATION NUMBER: US/10/422,536
; CURRENT FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: US 60/187,130
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 09/800,770
; PRIOR FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: US 10/232,758
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 136
; LENGTH: 2372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-422-536-136

Alignment Scores:
Pred. No.: 6.84e-130 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 15 Gaps: 1

US-10-057-510-4 (1-216) x US-10-422-536-136 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTACCACTTCCACAGATCCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuSer 40
DB 372 GCTTCGGAACAAGACACCTGCTGTAGACCAAGCCATGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 432 GTTGGTGACAAAAGACACTTACTATGAAAGAGTTCTTTTATCTTGCCTGAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 492 ATTATGACTAAACGATTATATGATGAGAACCAACATATTTATATCTTGCCTGAGTAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTTCTAGAGATTGTTTGGCGTCCCAAGCTTCTGTGAAAGAGCAGACAGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
DB 612 ACCATGATCTACAGNACTTGTGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisIleGluGlyCysSerAspGlnLysAspLeuValGln 139
DB 672 TCTGTAGTGAAGACAGGTGTCACCTTGAAGTGGAGTATCAAAAAGGACCTTGATCAA 731
QY 140 GluLeuGlnGluGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
DB 732 GAGCTTCAGGAAGAAACCTTCATCTCTACATTTGGTTTCTAGACCATCTCACTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
DB 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179

RESULT 12
US-10-005-344-324
; Sequence 324, Application US/10005344
; Publication No. US20030203862A1
; GENERAL INFORMATION:
; APPLICANT: Loren J. Miraglia
; APPLICANT: Pamela Nero
; APPLICANT: Mark J. Graham
; APPLICANT: Brett P. Monia
; APPLICANT: Erich Koller
; APPLICANT: Mingyi Chiang
; APPLICANT: Manoj Manoharan
; TITLE OF INVENTION: Antisense Modulation of mdm2 expression.
; FILE REFERENCE: ISPH-0622
; CURRENT APPLICATION NUMBER: US/10/005,344
; CURRENT FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: US 09/048,810
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: US 09/280,805
; PRIOR FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 379
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 324
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1470)
US-10-005-344-324

Alignment Scores:
Pred. No.: 6.17e-94 Length: 1470
Score: 805.00 Matches: 166
Percent Similarity: 84.72% Conservative: 17
Best Local Similarity: 76.85% Mismatches: 25
Query Match: 73.52% Indels: 8
DB: 15 Gaps: 3

US-10-057-510-4 (1-216) x US-10-005-344-324 (1-1470)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 1 ATGTGCAATACCAACATGCTGTGTCTACCGAGGCTGTGCAAGCACCTTCACAGATTCCA 60
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLysSer 40
DB 61 GCTTCGGAACAAGACACTGCTGTAGACCAACCAATTCCTTTTGAAGTTATTAAAGTCC 120
QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 121 GTTGGAGCGCAAAACGACACTTACACTATGAAAGAGATTATATTTATTTATTTGGCCAGTAT 180
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 181 ATTATGACTAAAGGTTTATATGACGAGAGCAGCAGCATTTGTGTATTGTTCAATGAT 240
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 241 CTTCTAGGAGATGTTGTTGGAGTCCCGAGTTTCTCTGTGAAAGAGCAGACAGAAATATAT 300
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
DB 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
```


Db 322 GTTCGACGCAAAACGACACTTACACTATGAAAGAGATTATATTTATATTGGCCAGTAT 381
Qy 60 IleMetThrLysArgLeuTyrAspGluLysGlnHisIleValTyrCysSerAsnAsp 79
Db 382 ATTATGACTAAGAGGTTATATGACGAGAACACACGACACATTTGTATTGTTCAAATGAT 441
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 442 CTCTAGGAGATGTGTTGGAGTCCCGAGTTTCTGTGAGGAGGACACAGGAAATATAT 501
Qy 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
Db 502 GCAATGATCTACAGAAATTTAGTGGCTGTAAAGTCAACAA-----GACTCTGGCACA 552
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 553 TCGCTGAGTGAGAGACACGCTCAGCCCTGAAGGTGGAGTGATCTGAAGGATCCTTTGCAA 612
Qy 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 613 CGGCCACACAGAGAAACCTTCATCTTCTGATTAAATTTCTAGACTGTCTACCTCATCT 672
Qy 160 ArgArgArgAlaIleSerGluThrGluLysAsnSerAspGluLeuSerGlyGluArgGln 179
Db 673 AGAAGGAGATCCATTAGTGACAGACAGAGAACACAGATGAGTACTCTGGGAGCGGCAC 732
Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 733 CGGAGCGCCGACAG-----TCCCTGTCTTTGATCCGAGCCTGGGTCTGTGT 780
Qy 200 ValIleArgGluIleCys-----CysGluArgSerSerSerGluSer 214
Db 781 GAGCTGAGGAGATGTGCGCGGCGCACGAGCAGCAGTAGCAGCAGCAGCAGCAGTCC 840
Qy 215 Thr 215
Db 841 ACA 843

RESULT 15
US-10-211-088-144
; Sequence 144, Application US/10211088
; Publication No US20030104479A1
; GENERAL INFORMATION:
; APPLICANT: Bright, Gary R.
; APPLICANT: Premkumar, D. David
; APPLICANT: Chen, Yih-Tai
; TITLE OF INVENTION: No. US20030104479A1el Fusion Proteins And Assays For Molecular Bi
; FILE REFERENCE: 01-1022-US
; CURRENT APPLICATION NUMBER: US/10/211,088
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/309,395
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/341,589
; PRIOR FILING DATE: 2001-12-13
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 144
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence encoding binding domain
US-10-211-088-144

Alignment Scores:
Pred. No.: 1,31e-28 Length: 176
Score: 290.50 Matches: 57
Percent Similarity: 98.28% Conservative: 0
Best Local Similarity: 98.28% Mismatches: 0
Query Match: 26.53% Indels: 1
DB: 14 Gaps: 1

US-10-057-510-4 (1-216) x US-10-211-088-144 (1-176)

Qy 42 GlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyrIle 60
Db 1 GGTGCACAAAAGACACTTACTATGAAAGAGTTCTTTTATCTTGGCCAGTATATT 60
Qy 61 MetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAspLeu 80
Db 61 ATGACTAAACGATTATATGATGAGAGCAACACATATTGTATATTGTTCAAATGATCTT 120
Qy 81 LeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIle 98
Db 121 CTAGGAGATTTGTTTGGCGTGCCCAAGCTTCTCTGTGAAAGAGCAGCAGGAAATA 174

Search completed: March 15, 2004, 00:28:55
Job time : 332.49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:33:39 / Search time 73.7379 Seconds
(without alignments)
924.245 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 1095

Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIREICRRSSSESTG 216

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1027.5	93.8	491	6 Q7YR28	Q7YR28 felis silve
2	1015.5	92.7	487	6 Q9GMZ6	Q9GMZ6 canis famil
3	940.5	85.9	195	4 Q96DS4	Q96DS4 homo sapien
4	860	78.5	243	4 Q8TE47	Q8TE47 homo sapien
5	829	75.7	446	4 Q8VJ1	Q8VJ1 homo sapien
6	805	73.5	489	11 Q91XK7	Q91XK7 m adult mal
7	774	70.7	436	4 Q8WJ2	Q8WJ2 homo sapien
8	742.5	67.8	166	4 Q8NDW2	Q8NDW2 homo sapien
9	733	66.9	426	6 Q9GK41	Q9GK41 canis famil
10	616.5	56.3	325	13 Q9PVL2	Q9PVL2 gallus gall
11	545.5	49.8	118	4 Q8VJ3	Q8VJ3 homo sapien
12	478.5	43.7	173	4 Q8TE45	Q8TE45 homo sapien
13	408.5	37.3	105	4 Q8NDW0	Q8NDW0 homo sapien
14	358.5	32.7	84	4 Q96DS2	Q96DS2 homo sapien
15	292.5	26.7	69	4 Q8GWA4	Q8GWA4 homo sapien
16	291.5	26.6	66	4 Q96DS3	Q96DS3 homo sapien

17	289	26.4	70	4	Q86WA3	Q86wa3 homo sapien
18	287.5	26.3	243	4	Q9H4C5	Q9h4c5 homo sapien
19	285.5	26.1	95	4	Q96DS1	Q96ds1 homo sapien
20	274.5	25.1	491	13	Q7ZUW7	Q7zuw7 brachydanio
21	274	25.0	489	11	Q9CYG1	Q9cyg1 mus musculus
22	272.5	24.9	490	11	Q99L86	Q99l86 mus musculus
23	266.5	24.3	159	4	Q96DS0	Q96ds0 homo sapien
24	266	24.3	70	4	Q8NDW1	Q8ndw1 homo sapien
25	265.5	24.2	475	13	Q7ZV13	Q7zv13 xenopus lae
26	262	23.9	60	4	Q96DS5	Q96ds5 homo sapien
27	262	23.9	130	4	Q9H4C3	Q9h4c3 homo sapien
28	261	23.8	65	4	Q86WA5	Q86wa5 homo sapien
29	242	22.1	50	4	Q8NDV9	Q8ndv9 homo sapien
30	203	18.5	54	4	Q86WA2	Q86wa2 homo sapien
31	134	12.2	98	4	Q9H4C2	Q9h4c2 homo sapien
32	110.5	10.1	151	4	Q9H4C4	Q9h4c4 homo sapien
33	103.5	9.5	851	5	O15892	O15892 tetrahymena
34	100	9.1	2518	5	O81EH2	O81eh2 plasmodium
35	99.5	9.1	1417	5	Q9VSE1	Q9vse1 drosophila
36	99.5	9.1	2253	13	P70012	P70012 xenopus lae
37	98.5	9.0	894	10	Q9FYB2	Q9fyb2 arabidopsis
38	97.5	8.9	650	13	O8JHG6	O8jhg6 brachydanio
39	95	8.7	507	5	O869T2	O869t2 dictyosteli
40	94.5	8.6	1650	5	O81ZT7	O81zt7 plasmodium
41	94	8.6	1107	4	Q9PZH7	Q9pzh7 homo sapien
42	93.5	8.5	2015	5	Q9USY1	Q9usy1 dictyosteli
43	93	8.5	302	12	Q8V7G7	Q8v7g7 tt virus. o
44	93	8.5	365	10	Q9SL41	Q9sl41 arabidopsis
45	93	8.5	530	4	Q9BWC1	Q9bwc1 homo sapien

ALIGNMENTS

RESULT 1

Q7YR28 ID Q7YR28 PRELIMINARY; PRT; 491 AA.

AC Q7YR28; 01-OCT-2003 (TREMBlrel. 25, Created)

DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DB Double minute 2 protein MDM2.

CN MDM2

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

OX NCBI_TaxID=9685;

RN [1]

RP SEQUENCE FROM N.A.

RA Miki R., Okuda M., Ma Z., Inokuma H., Onishi T.;

RT "Molecular cloning of feline mdm2 cDNA."

RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB099709; BAC78209.1; -

SQ SEQUENCE 491 AA; 55433 MW; D93E25D638E8934 CRC64;

Query Match

Best Local Similarity 93.8%; Score 1027.5; DB 6; Length 491;

Matches 205; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

Qy	1	MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLYL-QY 59
Db	1	MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLYLQY 60
Qy	60	IMTKRLYDEKQHHVVCSDNLLGLDGLFGVPSFVKHRKIYTMIRNLVWVNCQESSDST 119
Db	61	IMTKRLYDEKQHHVVCSDNLLGLDGLFGVPSFVKHRKIYTMIRNLVWVNCQESSDST 120
Qy	120	SVSENRCHLEGGSDQKDLVQELQEEKPSSHLVSRPSTSSRRRAISETENSDELSCERQ 179
Db	121	SVSENRCHLEGGSDQKDLVQELQEEKPSSDLVSRPSTSSRRRTISETEHSDELPCRQ 180
Qy	180	RKHKSDISLSFDESLAICVIREICRRSSSESTG 216


```
QY 180 RRRKSDSISLSPDESIALC 199
DB 174 KEESVESLPLNAIEPCVIC 193

RESULT 5
Q8WJ1 PRELIMINARY; PRT; 446 AA.
AC Q8WJ1
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MDM2 protein.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=21248713; PubMed=11351297;
RA Tamborini E., Della Torre G., Lavarino C., Azzarelli A.,
RA Carpinelli P., Pierotti M.A., Pilotti S.;
RT "Analysis of the molecular species generated by MDM2 gene
RT amplification in liposarcomas.";
RL Int. J. Cancer 92:790-796(2001).
DR EMBL; AF092845; AAL40180.1; -.
DR InterPro; IPR003121; SWIF.
DR InterPro; IPR001876; Znf RangDP.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF02201; SWIF; 1.
DR Pfam; PF00641; Zf-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS01359; ZF_RANBP2_2; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 446 AA; 4989 MW; 4B630B50750EADF CRC64;

Query Match 75.7%; Score 829; DB 4; Length 446;
Best Local Similarity 79.6%; Pred. No. 9.8e-64;
Matches 172; Conservative 0; Mismatches 0; Indels 44; Gaps 1;

QY 1 MONTNNSVPTDGAVTTSQIPASEQETLVPRPKLLKLSVGAQXDTYMKVELFYLQYI 60
DB 1 MONTNNSVPTDGAVTTSQIPASEQETLVPRPKLLKLSVGAQXDTYMKVE----- 52
QY 61 MTKRLYDEKQKHIVYCSNDLLGLFGVPFSVKEHRKIYTYMYRLVVVNQOESSDSGTS 120
DB 53 -----KIYTYMYRLVVVNQOESSDSGTS 76
QY 121 VSENCHLGGSDQKDLVQLOEKPSSSHLYSRPSTSRRAISGTETNSDELGERQOR 180
DB 77 VSENCHLGGSDQKDLVQLOEKPSSSHLYSRPSTSRRAISGTETNSDELGERQOR 136
QY 181 KKHKSDSISLSPDESIALCVIREICERSSSSSSTG 216
DB 137 KKHKSDSISLSPDESIALCVIREICERSSSSSSTG 172

RESULT 6
Q91XK7 PRELIMINARY; PRT; 489 AA.
AC Q91XK7
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Adult male lung cDNA, RIKEN full-length enriched library,
DE clone:1200011P22, full insert sequence (2 days neonate thymic
DE cells cDNA, RIKEN full-length enriched library, clone:E430022B10
DE product:transformed mouse 373 cell double minute 2, full insert
DE sequence) (Transformed mouse 373 cell double minute 2).
GN MDM2.
```

```
RC STRAIN=NOD: TISSUE=Thymus;
RA The FANTOM Consortium
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
[6]
RP SEQUENCE FROM N.A.
RP STRAIN=C57BL/6;
RP MEDLINE=22398257; PubMed=12477932;
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Iqbal N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Viallon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[9]
RP SEQUENCE FROM N.A.
RP STRAIN=C57BL/6;
RP Strausberg R.;
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; AK004719; BAB3502.1; -.
DR ENBL; AK088638; BAC40470.1; -.
DR ENBL; BC050902; AAH50902.1; -.
DR MGD; MGI:96952; Mdm2.
DR GO; GO:0005634; C:nucleus; IDA.
DR GO; GO:0005115; F:protein binding; IPI.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IDA.
DR GO; GO:0030163; P:protein catabolism; IDA.
DR GO; GO:0016567; P:protein ubiquitination; IDA.
DR GO; GO:0007089; P:start control point of mitotic cell cycle; IDA.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf_RanGDP.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF02201; SWIB; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS01099; ZF_RANBP2_2; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 489 AA; 54558 MW; 4AB489A82038DF4 CRC64;

Query Match 73.58; Score 805; DB 11; Length 489;
Best Local Similarity 76.98; Pred. No. 1.3e-61;
Matches 166; Conservative 17; Mismatches 25; Indels 8; Gaps 3;

QY 1 MCNTNMSVPTDGA VTTSGQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLYFYLYQY 59
DB 1 MCNTNMSVSTEGA AASTSQIPASEQETLVRPKPLLLKLLKSVGAQNDYTMKEIFVIGQY 60
QY 60 IMTKELYDEKQCHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLYVNVNQSSDSGT 119
DB 61 IMTKELYDEKQCHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLYVNVNQSSDSGT 117
QY 120 SVSENRCHEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISSETEENSDELGSRQ 179
DB 118 SLSESRQPEGGSDLKDLQAPPEKPSDDLISRLSTSSRRRSISTEENTDELGERH 177
QY 180 KRRHKSISLSFDESALCVIREICCRSSSSSTG 215

Query Match 70.78; Score 774; DB 4; Length 436;
Best Local Similarity 74.28; Pred. No. 5.7e-59;
Matches 161; Conservative 0; Mismatches 0; Indels 56; Gaps 2;

QY 1 MCNTNMSVPTDGA VTTSGQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLYFYLYQY 59
DB 1 MCNTNMSVPTDGA VTTSGQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLYFYLYQY 60
QY 60 IMTKELYDEKQCHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLYVNVNQSSDSGT 119
DB 61 IMTKELYDEKQCHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLYVNVNQSSDSGT 113
QY 120 SVSENRCHEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISSETEENSDELGSRQ 179
DB 114 -----EENSDELGSRQ 125
QY 180 KRRHKSISLSFDESALCVIREICCRSSSSSTG 216
DB 126 KRRHKSISLSFDESALCVIREICCRSSSSSTG 162

RESULT 8
Q8NDW2 PRELIMINARY; PRT; 166 AA.
ID Q8NDW2
AC Q8NDW2;
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE P53-binding protein.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

Db 178 KRRR-----SLSPDPSLGLCELRMCSSSSSSSS 209
RESULT 7
Q8WYJ2 PRELIMINARY; PRT; 436 AA.
ID Q8WYJ2
AC Q8WYJ2;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE MDM2 protein.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

MEDLINE=21248713; PubMed=11351297;
RA Tamborini E., Della Torre G., Lavarino C., Azzarelli A.,
RA Carpinelli P., Pierotti M.A., Pilotti S.;
RT "Analysis of the molecular species generated by MDM2 gene
RT amplification in liposarcomas."
RL Int. J. Cancer 92:790-796 (2001).
RL EMBL; AF092844; AAL40179.1; -.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf_RanGDP.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF02201; SWIB; 1.
DR Pfam; PF00641; ZF-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS01099; ZF_RANBP2_2; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
SQ SEQUENCE 436 AA; 49248 MW; 3C8F55E98BC4203A CRC64;
```


RA Bartel F., Pinkert D., Kappler M., Bache M., Schmidt H., Taubert H.;
RT "Alternatively and aberrantly spliced transcripts of the MDM2 mRNA
occur frequently in human soft tissue sarcoma and in multiple normal
tissues";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; AJ491698; CAD36959.1; -
DR InterPro; IPR003121; SWIB.
DR Pfam; PF02201; SWIB; 1.
SQ SEQUENCE 166 AA; 18900 MW; FA6B5BA18B85040D CRC64;

Query Match 67.8%; Score 742.5; DB 4; Length 166;
Best Local Similarity 96.1%; Pred. No. 9.9e-57;
Matches 147; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

QY 1 MONTNMSVPTDGAVTTSQIPASEQETLVPRKPKLLKLLKSVGAQKDTYTMKEVLYL-QY 59
DB 1 MONTNMSVPTDGAVTTSQIPASEQETLVPRKPKLLKLLKSVGAQKDTYTMKEVLYLQY 60

QY 60 IMTKRLYDEKQHIYVCSNDLLGDLFGVPSFSVKEHRKIYTYMYRNLVVNVNQESSDST 119
DB 61 IMTKRLYDEKQHIYVCSNDLLGDLFGVPSFSVKEHRKIYTYMYRNLVVNVNQESSDST 120

QY 120 SVSENRCHLEGGSDQKDLVQELQEEKPSSSHLV 152
DB 121 SVSENRCHLEGGSDQKDLVQELQEEKPSTKKRV 153

RESULT 9
Q9GK41
ID Q9GK41 PRELIMINARY; PRT; 426 AA.
AC Q9GK41;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MDM2 alpha.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Veldhoen N.J.; PubMed=10597303;
RX MEDLINE=20065171; Metcalfe S., Milner J.;
RT "A novel exon within the mdm2 gene modulates translation initiation in
vitro and disrupts the p53-binding domain of mdm2 protein.";
RL Oncogene 18:7026-7033 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Veldhoen N.J.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF322417; RAG4284.1; -
DR HSP; Q9UMT8; 1YCR.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf_RanGDP.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF02201; SWIB; 1.
DR Pfam; PF00641; zf-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS0199; ZF_RANBP2_2; 1.
DR PROSITE; PS0089; ZF_RING_2; 1.
SQ SEQUENCE 426 AA; 48064 MW; 40FAC2FB274AD4BF CRC64;

Query Match 66.9%; Score 733; DB 6; Length 426;
Best Local Similarity 92.3%; Pred. No. 2e-55;
Matches 144; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 61 MTKRLYDEKQHIYVCSNDLLGDLFGVPSFSVKEHRKIYTYMYRNLVVNVNQESSDST 120
DB 1 MTKRLYDEKQHIYVCSNDLLGDLFGVPSFSVKEHRKIYTYMYRNLVVNVNQESSDST 60

QY 121 VSENRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENDELUGBQR 180
DB 122 VSENRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENDELUGBQR 180

DB 61 VSENRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENDELUGBQR 120
QY 181 KRHKSDSISLSPDESALCVIREICCRSSSSSTG 216
DB 121 KRHKSDSISLSPDESALCVIREICCRSSSSSTG 156

RESULT 10
Q9PVL2
ID Q9PVL2 PRELIMINARY; PRT; 325 AA.
AC Q9PVL2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MDM2 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA LaFleur D.A.; Foster D.N.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF005045; AAF04192.1; -
DR HSP; Q9UMT8; 1YCR.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001876; Znf_RanGDP.
DR Pfam; PF02201; SWIB; 1.
DR Pfam; PF00641; zf-RanBP; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS0199; ZF_RANBP2_2; 1.
FT NON TER 325 325
SQ SEQUENCE 325 AA; 37205 MW; E3C8509CCF5FDIED CRC64;

Query Match 56.3%; Score 616.5; DB 13; Length 325;
Best Local Similarity 61.8%; Pred. No. 1.9e-45;
Matches 134; Conservative 35; Mismatches 39; Indels 9; Gaps 6;

QY 1 MONTNMSVPTDGAVTTSQIPASEQETLVPRKPKLLKLLKSVGAQKDTYTMKEVLYL-QY 59
DB 1 MONTNMSVPTDGAVTTSQIPASEQETLVPRKPKLLKLLKSVGAQKDTYTMKEVLYLQY 56

QY 60 IMTKRLYDEKQHIYVCSNDLLGDLFGVPSFSVKEHRKIYTYMYRNLVVNVNQESSDST 119
DB 57 IMTKRLYDEKQHIYVCSNDLLGDLFGVPSFSVKEHRKIYTYMYRNLVVNVNQESSDST 116

QY 120 SVSENRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENDELUGBQR 178
DB 117 PEMNPFGLKKMKPKSMQEL-EEKQTSNATSQP-TTSRRTHSESENSDDLLHSDR 174

QY 179 QRKRHKSDSISLSPDESALCVIREICCRSSSSSTG 215
DB 175 -RKRHKSDSISLSPDESALCVIREICCRSSSSSTG 210

RESULT 11
Q8WYJ3
ID Q8WYJ3 PRELIMINARY; PRT; 118 AA.
AC Q8WYJ3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MDM2 protein (Fragment).
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21248713; PubMed=11351297;
RA Tamborini E., Della Torre G., Lavarino C., Azzarelli A.,

```
RA Carpinelli P., Pierotti M.A., Pilotti S.;
RT "Analysis of the molecular species generated by MDM2 gene
RL Int. J. Cancer 92:790-796(2001).
DR EMBL: AF092843; AAL40178.1; -.
DR InterPro; IPR003121; SWIB.
DR Pfam; PF02201; SWIB; 1.
FT NON TER 118 118
SQ SEQUENCE 118 AA; 13536 MW; D7A4DBAA83D8841B CRC64;

Query Match 49.8%; Score 545.5; DB 4; Length 118;
Best Local Similarity 99.1%; Pred. No. 8.5e-40;
Matches 108; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVLRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVLRPKPLLLKLLKSVGAQKDTYTMKEVLFYLGQY 60

QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLV 108
Db 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLV 109

RESULT 12
Q8TE46 PRELIMINARY; PRT; 173 AA.
ID Q8TE46
AC Q8TE46;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MDM2 isoform N1.40 protein.
GN MDM2 ISOFORM N1_40.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Bartel F., Pinkert D., Kappler M., Bache M., Schmidt H., Taubert H.;
RT "Alternatively and Aberrantly Spliced Transcripts of the MDM2-mRNA
RT Occur Frequently in Human Soft Tissue Sarcomas and in Multiple Normal
RT Tissues.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ430614; CAD23252.1; -.
DR InterPro; IPR003121; SWIB.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF02201; SWIB; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
SQ SEQUENCE 173 AA; 19508 MW; AB26EECF8A003B261 CRC64;

Query Match 43.7%; Score 478.5; DB 4; Length 173;
Best Local Similarity 70.6%; Pred. No. 8.8e-34;
Matches 101; Conservative 9; Mismatches 12; Indels 21; Gaps 3;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVLRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVLRPKPLLLKLLKSVGAQKDTYTMKEVLFYLGQY 60

QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLV 108
Db 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLV 109

QY 120 SVS-----ENRCHLEG 130
Db 113 PNAIEPCVICQGRPKNGCIVHG 135

RESULT 13
Q8NDW0 PRELIMINARY; PRT; 105 AA.
ID Q8NDW0
AC Q8NDW0;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE P53-binding protein.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Bartel F., Pinkert D., Kappler M., Bache M., Schmidt H., Taubert H.;
RT "Alternatively and aberrantly spliced transcripts of the MDM2 mRNA
RT occur frequently in human soft tissue sarcoma and in multiple normal
RT tissues.";
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ491700; CAD36961.1; -.
DR InterPro; IPR003121; SWIB.
DR Pfam; PF02201; SWIB; 1.
SQ SEQUENCE 105 AA; 11940 MW; 289127D911672D63 CRC64;

Query Match 37.3%; Score 408.5; DB 4; Length 105;
Best Local Similarity 80.4%; Pred. No. 5.8e-28;
Matches 86; Conservative 5; Mismatches 11; Indels 5; Gaps 3;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVLRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVLRPKPLLLKLLKSVGAQKDTYTMKEVLFYLGQY 60

QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNL 106
Db 61 IMTKRLYDEKQOHIVYCSND--CANFLPLVDLSIR-----LYISNYITL 103

RESULT 14
Q96DS2 PRELIMINARY; PRT; 84 AA.
ID Q96DS2
AC Q96DS2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE MDM2 variant FB29.
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Rhabdomyosarcoma;
RA Bartel P., Taylor A.C., Taubert H., Harris L.C.;
RT "Novel mdm2 splice variants identified in pediatric rhabdomyosarcoma
RT tumors and cell lines.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF385325; AAL13245.1; -.
SQ SEQUENCE 84 AA; 8955 MW; 6BA9AE1E684F1D31 CRC64;

Query Match 32.7%; Score 358.5; DB 4; Length 84;
Best Local Similarity 49.1%; Pred. No. 9.8e-24;
Matches 82; Conservative 0; Mismatches 0; Indels 85; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVLRPKPLLLKLLKSVGAQKDTYTMKEVLFYQYI 60
Db 1 MCNTNMSVPTDGAVTTSQIPASEQETLVLRPKPLLLKLLKSVGAQKDTYTMKEVLFYQYI 60

QY 61 MTKRLYDEKQOHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVVVNQESSDSGTS 120
Db 28 -----ESSDSGTS 35

QY 121 VSENCHLEGGSDOKDLVQELQEBKPPSSSHLVSRPSTSSRRRAISET 167
Db 36 VSENCHLEGGSDOKDLVQELQEBKPPSSSHLVSRPSTSSRRRAISET 82
```

```

RESULT 15
Q86WA4
ID Q86WA4 PRELIMINARY; PRT; 69 AA.
AC Q86WA4;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE HDM2-HD3 protein.
GN HDM2.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Sturzenhocker B., Schlott T., Quentin T., Kube D., Jung W.,
RA Trumper L.,
RT "Abundant expression of spliced HDM2 in Hodgkin Lymphoma cells does
RT not interfere with p14ARF and p53 binding."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ50518; CAD79457.1; -.
KW Alternative splicing.
SQ SEQUENCE 69 AA; 7625 MW; CAD81EC6846E439A CRC64;

Query Match 26.7%; Score 292.5; DB 4; Length 69;
Best Local Similarity 98.4%; Pred. No. 4.2e-18;
Matches 60; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAOKDTVTMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAOKDTVTMKEVLFYLQY 60

QY 60 I 60
DB 61 I 61

Search completed: March 14, 2004, 21:47:45
Job time : 75.7379 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2004, 21:55:40 ; Search time 2676.17 Seconds
(without alignments)
2410.251 Million cell updates/sec

Title: US-10-057-510-4
Perfect score: 1095
Sequence: 1 MCNTNMSVPTGAVTTSQIP.....ALCVIREICRSESSSESTG 216

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_epool/US10057510/runat_09032004_162235_25991/app_query.fasta_1.654
-DB=EST -QFMT=fastcap -SUFFIX=std.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=apct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10057510@cgn_1_2921/runat_09032004_162235_25991 -NCPU=6 -ICPU=3
-NO WMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1023.5	93.5	970	12	BM479400	BM479400 AGENCOURT
2	813.5	74.3	709	10	BF057574	BF057574 7k46c07.x
3	811	74.1	1008	10	BF683210	BF683210 602139223
4	805	73.5	2999	11	AK088638	AK088638 Mus muscu
5	805	73.5	3028	11	AK004719	AK004719 Mus muscu
6	791	72.2	718	12	B1558930	B1558930 603240981
7	761	69.5	571	10	AW500516	AW500516 UI-HF-BNO
8	761	69.5	612	10	AW500514	AW500514 UI-HF-BNO
9	754	68.9	669	13	BY754572	BY754572 BY754572
10	742	67.8	573	10	AW176629	AW176629 IL4-CT007
11	738	67.4	617	10	AW822859	AW822859 uq21a07.y
12	728	66.5	687	13	BQ571946	BQ571946 UI-M-FCO-
13	703	64.2	754	13	BU708267	BU708267 UI-M-FCO-
14	672	61.4	669	10	AW231125	AW231125 uo39h02.y
15	664.5	60.7	482	14	R80343	R80343 yi56d02.r1
16	661	60.4	642	13	BY743471	BY743471 BY743471
17	661	60.4	663	13	BY745235	BY745235 BY745235
18	655	59.8	570	14	CD564111	CD564111 B0474A11-
19	650	59.4	656	10	BF548539	BF548539 UI-R-AO-a
20	645	58.9	650	13	BY743294	BY743294 BY743294
21	645	58.9	674	10	AK975448	AK975448 EST387557
22	642.5	58.7	778	10	BE900427	BE900427 601873652
23	631.5	57.7	821	13	BU371540	BU371540 603396558
24	628.5	57.4	518	9	AI902188	AI902188 IL-B7002-
25	625	57.1	524	14	CA871480	CA871480 K0910F02-
26	619	56.5	784	14	CD355548	CD355548 UI-M-FV0-
27	606	55.3	738	12	BQ977125	BQ977125 602846256
28	602	55.0	488	29	C6303339	C6303339 OST345509
29	598	54.6	702	10	BE535040	BE535040 601233236
30	598	54.6	3007	11	AK004881	AK004881 Mus muscu
31	596	54.4	663	13	BY746837	BY746837 BY746837
32	581	53.1	468	14	CD562678	CD562678 B0451D06-
33	581	53.1	612	14	CA890706	CA890706 B0161A01-
34	578	52.8	468	14	CD549480	CD549480 B0299B01-
35	578	52.8	602	14	CA890782	CA890782 B0151E05-
36	569	52.0	638	14	CF903758	CF903758 A0413E01-
37	567	51.8	526	10	BF220646	BF220646 uz55b12.y
38	562.5	51.4	964	10	BE300019	BE300019 600944138
39	561	51.2	674	10	BF161075	BF161075 601768015
40	548	50.0	446	14	CF165114	CF165114 B0752C07-
41	543	49.6	451	14	CF162052	CF162052 B0705G01-
42	541	49.4	621	13	BU610936	BU610936 UI-M-FCO-
43	536.5	49.0	393	9	AL704062	AL704062 DRF2B686N
44	534	48.8	443	29	CG520319	CG520319 OST85447
45	533	48.7	576	13	BY705086	BY705086 BY705086

ALIGNMENTS

RESULT 1
BM479400
LOCUS
DEFINITION BM479400 6418503 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5502336
5' mRNA sequence.
ACCESSION BM479400.1 GI:18528442
VERSION BM479400
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 970)

QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysLeuLysSer 40
 Db 266 GCTTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAGTCT 325
 QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
 Db 326 GTTGGTGCAACAAAAGACACTTATCTATGTAAGAGAGGTTCTTTTATCTTGGCCAGTAT 385
 QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisLeValTyrCysSerLeuAsp 79
 Db 386 ATTATGACTAAAGATATATGATGAGAGCAACACATATGTTATTTGTCAAATGAT 445
 QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysLeuTyr 99
 Db 446 CTTCTAGGAGATTGTTGGCGTCCCAAGCTTCTGTGAAGAGACACAGGANAATATAT 505
 QY 100 ThrMetIleTyrArgLeuValValValAsnGlnGlnLysSerSerAspSerGlyThr 119
 Db 506 ACCATGATCACAGGAACCTTGGTAGTACATCAGCAGGAATCATCGGACTCAGGTACA 565
 QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
 Db 566 TCTGTGAGTGAACAAGGTGTCACTTGAAGTGGGAGTGATCANAGGACCTTGTACAA 625
 QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
 Db 626 GAGCTTCATGAAGAGAAACCTTCATCTCACATTTGGTTTCTAGACCATCTACCTCATCT 685
 QY 160 ArgArgArgAlaIleSerGluThr 167
 Db 686 AGAAGAGAGCAATTGTGAGACC 709

RESULT 3

BF683210

LOCUS

DEFINITION 602139223F1 NIH_MSC_46 Homo sapiens cDNA clone IMAGE:4298285 5',
 mRNA sequence.

ACCESSION

BF683210

VERSION

BF683210.1

SOURCE

Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NIH-MGC http://imgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: ARCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: LCM1183 row: m column: 06
 High quality sequence stop: 697.
 Location/Qualifiers

1. .1008

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4298285"

/tissue_type="leiomyosarcoma cell line"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 46"

/note="Organ: uterus; Vector: pOTB7; Site: 1: XhoI; Site: 2:

EcoRI; cDNA made by oligo-dT priming. Directionally cloned

into EcoRI/XhoI sites using the following 5' adaptor:

GCCACGAGG). Size-selected >500bp for average insert size

1.8kb. Library constructed by Ling Hong in the laboratory

source

FEATURES

AK088638
 Mus musculus 2 days neonate thymus cells cDNA, RIKEN
 full-length enriched library, clone:E43022B10 product:transformed
 mouse 3T3 cell double minute 2, full insert sequence.

AK088638

AK088638.1

GI:26353679

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

of Gerald M. Rubin (University of California, Berkeley)
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript
 II RT (Life Technologies). Note: this is a NIH_MGC
 Library."

ORIGIN

Alignment Scores:

Pred. No.: 3,87e-78 Length: 1008
 Score: 811.00 Matches: 177
 Percent Similarity: 81.00% Conservative: 2
 Best local Similarity: 80.09% Mismatches: 6
 Query Match: 74.06% Indels: 36
 DB: 10 Gaps: 2

US-10-057-510-4 (1-216) x BF683210 (1-1008)

QY 1 MetCysLeuThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
 Db 56 ATGTGCAATACCAACATGCTGTACTACTGATGCTGTGTAAACCACTCACAGATTCCA 115
 QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysLeuLysSer 40
 Db 116 CTTTCGGNACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAGTCT 175
 QY 41 ValGlyAlaGlnLysAspThrThrMetLysGlu----- 52
 Db 176 GTTGGTGCAACAAAAGACACTTATCTATGAAAGAGAGATGGAGTTTCACTATCTTGCCC 235
 QY 52 ----- 52
 Db 236 AGGCTGTGCTGGAACTCCTGGGCTCAAGGGATCTGTACTCTGGCCCTCTCTAAAGTGCTA 295
 QY 53 -----ValLeuPheTyrLeu---GlnTyrIleMetThrLysArgLeuTyrAspGlu 68
 Db 296 GATTTCACAGGTTCTTTTATCTTGGCCAGTATATATGACTAAACCATTAATGATGAG 355
 QY 69 LysGlnGlnHisIleValTyrCysSerAsnAspLeuLeuGlyAspLeuPheGlyValPro 88
 Db 356 AAGCAACAACATATGTATATTTTCAATGATCTTCTAGGAGATTTGTTGGCGTCCA 415
 QY 89 SerPheSerValLysGluHisArgLysIleTyrThrMetIleTyrArgAsnLeuVal 108
 Db 416 AGCTTCTCTGTGAAGAGCACAGCAATATATACCATGATCTACAGAACTTGTAGTA 475
 QY 109 ValAsnGlnGlnLysSerSerAspSerGlyThrSerValSerGluAsnArgCysHisLeu 128
 Db 476 GTCAATCAGCAGGAATCATCGGACTCAGGTACATCTGTGAGTGAGAACAGGTGCACCTT 535
 QY 129 GluGlyGlySerAspGlnLysAspLeuValGln-GluLeuGlnGlu-GluLysProSer 148
 Db 536 GAAGTCGGAGTGATCAAGAGACCTTGTACAATCAGCTTCAGGAACGAGAAACCTTCAT 595
 QY 148 erSerHisLeuValSerArgProSerThrSerArgArgAlaIleSerGluThrG 168
 Db 596 CTTTCACATTAGGTTCTTAGACCATCTACCTCATCTAGAGAGAGAGCAATTAAGTACAG 655
 QY 168 luGluAsnSer-AspGlu-LeuSerGlyGluArgGlnArgLysArgHisLysSer 185
 Db 656 AAGAAACATCAGCATGACATTATCTGTGTGAACGACAAAGAAAGGACACACAAAGTCG 710

RESULT 4

AK088638

LOCUS

DEFINITION

AK088638
 Mus musculus 2 days neonate thymus cells cDNA, RIKEN
 full-length enriched library, clone:E43022B10 product:transformed
 mouse 3T3 cell double minute 2, full insert sequence.

ACCESSION

AK088638

VERSION

AK088638.1

HTC; CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.


```
Db 254 ATGTGCAATACCAACATGCTGTGTCTACCGAGGGTGTCTGCAAGCAGCCTCACAGATTCCA 313
Qy 21 AlaserGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLysSer 40
Db 314 GCTTCGGAACAGAGACTCTGGTTAGACCAAAACCAATCTCTTTGAAGTTGTAAAGTCC 373
Qy 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu---GlnTy 59
Db 374 GTTGGAGCGCAAAACACACTTACATCATGAAGAGATTATATTATATGCGCAGTAT 433
Qy 60 IleMetThrLysArgLeuTyAspGluLysGlnGlnHisIleValTyCysSerAsnAsp 79
Db 434 ATTATGACTAAGAGGTTATATGACGAGAAACGACGACACATTTGTATTGTTCAATGAT 493
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTy 99
Db 494 CTCTAGGAGATGTGTTGGAGTCCCGAGTTCTCTGTGAAGAGCAGCAGGAATATAT 553
Qy 100 ThrMetIleTyArgAsnLeuValValAsnGlnGlnLysSerAspSerGlyThr 119
Db 554 GCAATGATCTACAGAAATTTAGTGGCTGTAAAGTCAGCAA-----GACTCTGCACA 604
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139
Db 605 TCCTGAGTGAGAGCAGACGCTCAGCTGAGGTTGGAGGTGATCTGAAGGATCCTTTGCAA 664
Qy 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 665 GCSCCACCAGAGAGAACTTCATCTCTGATTTAATTTCTAGACTGTCTACCTCATCT 724
Qy 160 ArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 725 AGAAGAGATCCATTATGTAGACAGAGAGAAACACAGATGAGCTACCTCGGGAGCGGCAC 784
Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGlnSerLeuAlaLeuCys 199
Db 785 CGAAGCGCGCAGG-----TCCTGTCTCTTTGATCCCGAGCCTGGGTCTGTCT 832
Qy 200 ValIleArgGluLeuCysCysGluArgSerSerSerSerGluSerThr 215
Db 833 GAGCTGAGGAGATGTGTCAGCGCGGCGAGCAGCAGCAGTAGCAGCAGC 880
```

```
RESULT 6
BI558930 718 bp mRNA linear EST 05-SEP-2001
LOCUS 603240981F1 NCI_CGAP_Mam4 Mus musculus cDNA clone IMAGE:5293725 5',
DEFINITION mRNA sequence.
ACCESSION BI558930
VERSION BI558930.1 GI:15446244
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 718)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-1@mail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Priscilla Furch
Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM11743 row: a column: 22
High quality sequence stop: 711.
FEATURES
Location/Qualifiers
1..718
source
```

```
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NMRI"
/db_xref="taxon:10090"
/clone="IMAGE:5293725"
/tissue_type="tumor, gross tissue"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI CGAP Mam4"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: Sali;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Priscilla Furch,
NIH Reference for transgenic model: Li et al., Cell Growth
and Differentiation 7, 3-11 (1996)."
```

ORIGIN

```
Alignment Scores: 3.61e-76 Length: 718
Pred. No.: 791.00 Matches: 164
Score: 791.00 Conservative: 17
Percent Similarity: 84.58% Mismatches: 25
Best Local Similarity: 76.64% Indels: 8
Query Match: 72.24% Gaps: 3
DB: 12

US-10-057-510-4 (1-216) x BI558930 (1-718)
Qy 3 AsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIleProAlaSer 22
Db 3 AATACCAACATGTCTGTGTCTACCGAGGTGCTGCAAGCAGCCTCACAGATTCAGCTTCG 62
Qy 23 GluGlnGluThrLeuValArgProLysProLeuLeuLeuLysSerValGly 42
Db 63 GAAACAGAGACTCTGGTTAGACCAAAACCAATGCTTTTGAAGTTGTTAAAGTCCGTGGA 122
Qy 43 AlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu---GlnTyIleMet 61
Db 123 CGCAGAAACGACATTTACACTATGAAGAGATTATATTTATATTTGCGCAGTATATTATG 182
Qy 62 ThrLysArgLeuTyAspGluLysGlnGlnHisIleValTyCysSerAsnAspLeuLeu 81
Db 183 ACTAAGAGGTTATATGACGAGAACGACAGCAGCATTTGTATTGTTCAATGATCTCTCTA 242
Qy 82 GlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyThrMet 101
Db 243 CGAGATGTTGTTGAGTCCCGAGTTTCTGTGAGAGCAGCAGCAAAATATATGCAATG 302
Qy 102 IleTyArgAsnLeuValValAsnGlnGlnLysSerAspSerGlyThrSerVal 121
Db 303 ATCTACAGAAATTTAGTGGCTGTAAAGTCAGCAA-----GACTCTGGCAGCATCGCTG 353
Qy 122 SerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGlnGluLeu 141
Db 354 AGTGAGAGCAGACGCTCAGCTGAGAGTGGGAGTGATCTGAAGGATCTTTGCAAGCGCCA 413
Qy 142 GlnGluGluLysProSerSerHisLeuValSerArgProSerThrSerSerArgArg 161
Db 414 CCAGAGAGAAACCTTCATCTCTGATTTAATTTCTAGACTGTCTACTCATCTAGAAAG 473
Qy 162 ArgAlaIleSerGluThrGluLysSerAspGlnLysSerGlyGluArgGlnArgLys 181
Db 474 AGATCCATTAGTAGACAGAGAGACAGATAGCTACTCTGGGAGCGGACCCGGAAG 533
Qy 182 ArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCysValIle 201
Db 534 CGCCGCGCAGG-----TCCTGTCTCTTTGATCCGAGCTGGGTCTGTGTGAGCTG 581
Qy 202 ArgGluIleCysCysGluArgSerSerSerSerGluSerThr 215
Db 582 AGGAGAGATGTCAGCGCGGCGCAGCAGCAGCAGTAGTAGCAGCAGC 623
```

RESULT 7
AWS00516

LOCUS AW500516 571 bp mRNA linear EST 01-MAR-2000
 DEFINITION UI-HF-BNO-akj-b-08-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone
 IMAGE:3077318 5', mRNA sequence.
 ACCESSION AW500516
 VERSION AW500516.1 GI:7113204
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 571)
 NIH-MGC http://mhc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Eco RI site shown at the beginning of the sequence.
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M.B. Soares Lab
 CDNA Library Arrayed by: M.B. Soares Lab
 DNA Sequencing by: M.B. Soares Lab
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward.

FEATURES
 source
 1..571
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3077318"
 /tissue_type="lymph"
 /cell_type="germinal center B cells"
 /cell_line="MGC85"
 /lab_host="DH10B (LTI)"
 /clone_lib="NIH_MGC_50"
 /note="Vector: p77T3-Pac; Site_1: NotI; Site_2: Eco RI;
 Constructed from size fractionated cytoplasmic mRNA
 (3.5-4.4kb). Directionally cloned. Cells provided by
 Louis M. Staudt, Ph.D. Library preparation by Maria de
 Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN
 Alignment Scores:
 Pred. No.: 5e-73 Length: 571
 Score: 761.00 Matches: 150
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 69.50% Indels: 0
 DB: 10 Gaps: 0

US-10-057-510-4 (1-216) x AW500516 (1-571)

Qy 67 AspGluLysGlnGlnHisIleValTyrCysSerAsnAspLeuLeuGluYAspLeuPheGly 86
 Db 7 GATGAGAGCAACACATATTGTATATTGTCATATGATCTCTAGGAGATTGTTGGC 66

Qy 87 ValProSerPheSerValLysGluHisArgLysIleTyrThrMetIleTyrArgAsnLeu 106
 Db 67 GTGCAAGCTTCTCTGTGAAGAGACACAGGAAAATATATACCATGATCTACAGGAACCTG 126

Qy 107 ValValLysGlnGlnGluSerSerAspSerGlyThrSerValSerGluAsnArgCys 126
 Db 127 GTAGTAGTCATACAGCAGGAATCATCGGACTCAGGTACATCTGTGATGAGAACAGGTGT 186

Qy 127 HisLeuGluGlyGlySerAspGlnLysAspLeuValGlnGlnLeuGlnGluLysPro 146
 Db 187 CACCTTGAGGTGGAGTGATCAAAAGACCTTGTCACAGAGCTTCAGGAAGAAACCT 246

Qy 147 SerSerSerHisLeuValSerArgProSerThrSerSerArgArgAlaIleSerGlu 166
 Db 247 TCATCTTCACATTGGTTTCTAGACCATCTACCTCATCTAGAAGGAGCAATTAGTGAG 306

Qy 167 ThrGluGluAsnSerAspGluLeuSerGlyGluArgGlnArgLysHisLysSerAsp 186
 Db 307 ACAGAGAGAAATTCAGATGAATTACTGCTGACACACAGAAAGCCCAAAATCTGAT 366

Qy 187 SerIleSerLeuSerPheAspGluSerLeuAlaLeuCysValIleArgGluIleCysCys 206
 Db 367 AGTATTTCCTTTCTCTTGTGATGAAGCCCTGCTGTGTGTAATAAGGAGATATGTTGT 426

Qy 207 GluArgSerSerSerSerGluSerThrGly 216
 Db 427 GAAAGAAGCAGTAGCAGTGAATCTACAGG 456

RESULT 8
 LOCUS AW500514
 DEFINITION UI-HF-BNO-akj-b-06-0-UI.r1 NIH_MGC_50 Homo sapiens cDNA clone
 IMAGE:3077314 5', mRNA sequence.
 ACCESSION AW500514
 VERSION AW500514.1 GI:7113200
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 612)
 NIH-MGC http://mhc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Eco RI site shown at the beginning of the sequence.
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
 cDNA Library Preparation: M.B. Soares Lab
 CDNA Library Arrayed by: M.B. Soares Lab
 DNA Sequencing by: M.B. Soares Lab
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: M13 Forward.

FEATURES
 source
 1..612
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3077314"
 /tissue_type="lymph"
 /cell_type="germinal center B cells"
 /cell_line="MGC85"
 /lab_host="DH10B (LTI)"
 /clone_lib="NIH_MGC_50"
 /note="Vector: p77T3-Pac; Site_1: NotI; Site_2: Eco RI;
 Constructed from size fractionated cytoplasmic mRNA
 (3.5-4.4kb). Directionally cloned. Cells provided by
 Louis M. Staudt, Ph.D. Library preparation by Maria de
 Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."

ORIGIN
 Alignment Scores:
 Pred. No.: 5.56e-73 Length: 612
 Score: 761.00 Matches: 150
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 69.50% Indels: 0
 DB: 10 Gaps: 0

US-10-057-510-4 (1-216) x AW500514 (1-612)

Qy 67 AspGluLysGlnGlnHisIleValTyrCysSerAsnAspLeuLeuGluYAspLeuPheGly 86
 Db 7 GATGAGAGCAACACATATTGTATATTGTCATATGATCTCTAGGAGATTGTTGGC 66

Qy 87 ValProSerPheSerValLysGluHisArgLysIleTyrThrMetIleTyrArgAsnLeu 106

QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValValGluHisArgLysIleTyr 99
 DB 323 CTCCTAGGAGATGTTTGGATCCGAGTTTCTCTGTGAAGGACACAGAAATATAT 382

QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerAspSerGlyThr 119
 DB 383 GCAATGATCTACAGAAATTTAGTGGCTGTAACTCAGCAA-----GACTCTGGCACA 433

QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
 DB 434 TCCTCTAGGAGACAGACGCTCAGCTGAAGTGGAGTGATCTGAAGATCCCTTTGCAA 493

QY 140 GluLeuGlnGluGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
 DB 494 GCGCACCAAGAGAGAACTTCATCTCTGATTTAAATCTAGACTGTCTACCTCACT 553

QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
 DB 554 AGAAGGAGATCCCATTTAGTGAGACAGAGAGAACACAGATGACTACCTGGGAGCGGCA 613

QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
 DB 614 CCGGAGCGCGC-----ANGTCCCTGTCTTTGATCCGAGCGCTGGTCTGTGT 661

RESULT 10
 AW176629
 LOCUS IL4-CT0079-200899-001-B01 CT0079 Homo sapiens cDNA, mRNA sequence.
 ACCESSION AW176629
 VERSION AW176629.1 GI:6442666
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 573)
 HCGP <http://www.ludwig.org.br/ORESTES>.
 TITLE The FAPESP/LICR Human Cancer Genome Project
 JOURNAL Unpublished (1999)
 COMMENT Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL4&t2=IL4-CT0079-200899-001-B01&t3=1999-08-20&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence stop: 472.
 Location/Qualifiers
 1..573
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="CT0079"
 /note="Organ: colon; Vector: pUC18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
 Alignment Scores:
 Pred. No.: 6,138-71 Length: 573
 Score: 742.00 Matches: 147

Percent Similarity: 99.32% Conservative: 0
 Best Local Similarity: 99.32% Mismatches: 1
 Query Match: 67.76% Indels: 0
 DB: 10 Gaps: 0

US-10-057-510-4 (1-216) x AW176629 (1-573)

QY 69 LysGlnGlnHisIleValTyrCysSerAsnAspLeuLeuGlyAspLeuPheGlyValPro 88
 DB 2 AAGCAACAATATTTGTTATTTCTTCAATGATCTCTAGGAGATTTGTTGGGTGCCA 61

QY 89 SerPheSerValLysGluHisArgLysIleTyrThrMetIleTyrArgAsnLeuVal 108
 DB 62 AGCTTCTCTGTGAAGACACAGAAATATATACCATGATCTACAGAACTTGGTAGTA 121

QY 109 ValAsnGlnGlnGluSerSerAspSerGlyThrSerValSerGluAsnArgCysHisLeu 128
 DB 122 GTCATCAGCAGGAATCATCGGACTCAGGTACATCTGTGAGTGAGAACAGGTTCACCTT 181

QY 129 GluGlyGlySerAspGlnLysAspLeuValGlnGluLeuGlnGluLysProSerSer 148
 DB 182 GAAGTGGGAGTCAATCAAAAGGACCTTGTACAAGAGCTTCAGGAAGAGAAACCTTCATCT 241

QY 149 SerHisLeuValSerArgProSerThrSerArgArgArgAlaIleSerGluThrGlu 168
 DB 242 TCACATTTGTTTCTAGACCATCTACCTCATCTAGAAGGAGAGCAATTAGTGACAGAA 301

QY 169 GluAsnSerAspGluLeuSerGlyGluArgGlnArgLysArgHisLysSerAspSerIle 188
 DB 302 GAAATTCAGATCAATTTATCTGTGTGAACGACAAAGAAACGCCACATATCTGATAGTATT 361

QY 189 SerLeuSerPheAspGluSerLeuAlaLeuCysValIleArgGluLysCysGluArg 208
 DB 362 TCCCTTTTCTTGTATGAAGCTGGCTGTGTGTATTAAGAGAGATATGTTGTGAAGA 421

QY 209 SerSerSerSerGluSerThrGly 216
 DB 422 AGCAGTAGCAGTGATCTACAGG 445

AW822859 617 bp mRNA linear EST 17-MAY-2000
 uc21a07.y1 Ren Stubbs mouse thymus Mus musculus cDNA clone
 IMAGE:2803092.5' similar to gb.X58876 Murine mdm2 mRNA for mdm2
 protein (MOUSE);, mRNA sequence.
 ACCESSION AW822859
 VERSION AW822859.1 GI:7915936
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 617)
 Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
 Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
 Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
 Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
 Waterston,R. and Wilson,R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Other ESTs: uc21a07.x1
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseset@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:1042616
 Seq primer: Primer name ambiguous
 High quality sequence stop: 473.
 Location/Qualifiers

FEATURES

```

source
1. .617
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C3H"
/db_xref="taxon:10090"
/clone="IMAGE:2803092"
/sex="mixed"
/dev_stage="3 weeks"
/lab_host="DH10B"
/clone_lib="Ren Stubbs mouse thymus"
/notes="Organ: thymus; Vector: p773D-Pac; Site 1: NotI;
Site 2: PacI; 1st strand cDNA was primed with an oligo(dT)
primer; double-stranded cDNA was ligated using 5' linker
gagcggtat and 3' linker actggaagcttaatt. Library is
size-selected >2.5 kb and average insert size is 3.5 kb.
Clones were arrayed from primary plating; non-amplified.
Library constructed by X. Ren and L. Stubbs (Lawrence
Livermore National Laboratory and DOE Joint Genome
Institute, 7000 East Ave, L-453, Livermore, CA 94550)."

ORIGIN
Alignment Scores:
Pred. No.: 1.89e-70 Length: 617
Score: 738.00 Matches: 154
Percent Similarity: 84.42% Conservative: 14
Best Local Similarity: 77.39% Mismatches: 23
Query Match: 67.40% Indels: 8
DB: 10 Gaps: 3

US-10-057-510-4 (1-216) x AW822859 (1-617)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 42 ATGTGCAATACCAACATGCTGTGTCTACCGGGTCTGCAAGACCTTCACAGATCCA 101
QY 21 AlaSerGluGlnThrLeuValArgProLysProLeuLeuLeuLeuLeuLysSer 40
Db 102 GCTTCGGAACAGAGCTCTGGTTAGACCAAAACCATTCCTTTGAAGTTGTTAAGTCC 161
QY 41 ValGlyAlaGlnLysAspThrThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 162 GTTGGAGCGCAAAACGACACTTACATGATGAAGAGATTATATTTTATATTCGCCAGTAT 221
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisLeuValTyrCysSerAsnAsp 79
Db 222 ATTATGACTAAGAGGTTATATGACGAGACGACACACATGCTGTATGTTCAATGAT 281
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 282 CTCTAGAGATGTTGTTGGAGTCCCGAGTTTCTGTGTAAGGAGCACAGGAAATATAT 341
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
Db 342 GCATGATCTACAGAAATTTAGTGGCGGTAAAGTCAGCA-----GACTCTGGCACA 392
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 393 TCGTGAGTCAGACGACGACGCTCAGCTGAAGGTGGGAGTGATCTGAAGGATCCTTTGCAA 452
QY 140 GluLeuGlnGluLysProSerSerSerHisLeuValSerArgProSerThrSerSer 159
Db 453 GCGCCACCAAGAGAGAACTTCATCTCTGATTTAATTTTATAGACTGTCTACCTCATCT 512
QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 513 AGAAGGATGCCATTAGTGAGACAGAGAGAAACACAGATGAGCTACTCGGGAGCGGCAC 572
QY 180 ArgGlyArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeu 198
Db 573 CGGAAGCGCCCGCAG-----GTCTCTCTCTGTGATCCGAGCCTGGGCTGTG 617

RESULT 12
BQ571946
LOCUS BQ571946 687 bp mRNA linear EST 19-JUN-2002
DEFINITION UI-M-FCO-byc-p-22-0-UI.r1 NIH BMAP_FCO Mus musculus cDNA clone
IMAGE:5716485 5', mRNA sequence.
ACCESSION BQ571946
VERSION BQ571946.1 GI:21475263
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 687)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: PYX-5.
Location/Qualifiers
1. .687
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:5716485"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_FCO"
/notes="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with EcoR I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA
tail, is TCAGAGAGCC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): 'Gene Discovery in the Developing Mouse Nervous
System', supported by National Institutes of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
ORIGIN
Alignment Scores:
Pred. No.: 2.79e-69 Length: 687
Score: 728.00 Matches: 148
Percent Similarity: 87.98% Conservative: 13
Best Local Similarity: 80.87% Mismatches: 18
Query Match: 66.48% Indels: 4
DB: 13 Gaps: 2

US-10-057-510-4 (1-216) x BQ571946 (1-687)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 147 ATGTGCAATACCAACATGCTGTGTCTACCGGGTCTGCAAGACCTTCACAGATCCA 206
QY 21 AlaSerGluGlnThrLeuValArgProLysProLeuLeuLeuLeuLysSer 40
Db 207 GCTTCGGAACAGAGACTCTGGTTAGACCAAAACCATTCCTTTGAAGTTGTTAAGTCC 266
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59

```

```
Db 267 GTTGGAGCGCAAAACGACACTTACCTATGAAAGAGATTATATTTATTTGGCCAGTAT 326
Qy 60 IleMetThrLysArgLeuTyrAspGluLysGlnHisLeuValTyrCysSerAsnAsp 79
Db 327 ATTATGACTAAGAGGTATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 386
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 387 CTCCTAGAGATGTTTGGAGTCCCGAGTTCTCTGCTGAGGAGACACAGAAATATAT 446
Qy 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnSerAspSerGlyThr 119
Db 447 GCAATGATCTACAGAAATTTAGTGGCTGTAAGTCAGCAA-----GACTCTGGCACA 497
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 498 TCGTGTAGTACAGACAGACGTCACGCTGAGCTGGAGTGATCTGAAGGATCTTTGCCAA 557
Qy 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 558 CGCGCACAGAGAGAAACCTTCATCTCTGATTATTTCTAGACTGTCTACCTCATCT 617
Qy 160 ArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 618 AGAAGGAGATCCATTAGTGACAGAGAGAGAGAGAGATGAGCTACCTGNGGAGCGGCAC 677
Qy 180 ArgLysArg 182
Db 678 CGGAAGCGC 686
```

```
RESULT 13
BU708267
LOCUS
DEFINITION BU708267 754 bp mRNA linear EST 15-JUL-2003
IMAGE:6415937 5', mRNA sequence.
ACCESSION BU708267
VERSION BU708267.1 GI:23640562
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 754)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
The following repetitive elements were found in this cDNA
sequence: 65-95, >GC_rich#Low_complexity (matched complement)
seg primer: pYX-5.
FEATURES
source
1..754
location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6415937"
/tissue_type="whole brain"
/dev_stage="embryo 12.5 dpc"
/lab_host="DH10B (TI phage resistant)"
/clone_lib="NIH_BMAP_FCO"
/note="Organ: brain; Vector: pYX-Asc; Site_1: EcoR I;
```

Site 2: Not 1; The library was constructed according to Bonafide, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I, and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAAGAGGC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

ORIGIN

```
Alignment Scores: 1.79e-66 Length: 754
Pred. No.: 703.00 Matches: 148
Score: 703.00
Percent Similarity: 87.43% Conservative: 12
Best Local Similarity: 80.87% Mismatches: 19
Query Match: 64.20% Indels: 5
DB: 13 Gaps: 2
US-10-057-510-4 (1-216) x BU708267 (1-754)
Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 215 ATGTGCATPACCAACATGCTGTGTCTACCGAGGGTGTGCAAGCACCTCAGATTCCA 274
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysSer 40
Db 275 GCTTCGGACACAGAGACTCTGCTTAGACCAACCACTTCTTTGAAGTTGTAAGTCC 334
Qy 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 335 GTTGGAGCGCAAAACGACACTTACACTATGAAGAGATTATATTTATTTGGCCAGTAT 394
Qy 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisLeuValTyrCysSerAsnAsp 79
Db 395 ATTATGACTAAGAGGTTATATGACGAGAGCAGCAGCAGCATTTGTGTTTCAATGAT 454
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 455 CTCCTAGGAGATGTGTTTGGAGTCCCGAGTTCTCTGTAAGAGACACAGGAAATATAT 514
Qy 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnSerSerAspSerGlyThr 119
Db 515 GCAATGATCTACAGAAATTTAGTGGCTGTAAGTCAGCAA-----GACTCTGGCACA 565
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 566 TCGCTGAGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 625
Qy 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 626 GCGCCACAGAGAGAAACCTTCTCTGATTATTTCTAGACTGTCTACCTCATCT 685
Qy 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 686 AGAAGGAGATTCATTAGTGAGAGAGAGAGAGAC-ACAGATGAGCTACCTGGGAGCGGCAC 744
Qy 180 ArgLysArg 182
Db 745 CGGAAGCGC 753
RESULT 14
BU708267
LOCUS
DEFINITION BU708267 669 bp mRNA linear EST 10-DEC-1999
IMAGE:6415937 5', mRNA sequence.
Accession: U03902.y1 NCI CGAP Lu29 Mus musculus cDNA clone IMAGE:264947 5',
similar to gb:XS8876 Murine mdm2 mRNA for mdm2 protein (MOUSE);,
mRNA sequence.
Accession: AW231125
```

```

VERSION      AW231125.1  GI:6560421
KEYWORDS     Mus musculus (house mouse)
SOURCE       Mus musculus
ORGANISM     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 669)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE        National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
             Tumor Gene Index
JOURNAL      Unpublished (1997)
COMMENT      Other ESTs: uc039u02.x1
             Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-re@mail.nih.gov
             Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium, Inc.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             www-bio.llnl.gov/bbrp/image/image.html

MGI:1025399
Seq primer: -40RP from Gibco
High quality sequence stop: 411.
FEATURES     source
             1..669
             /organism="Mus musculus"
             /mol_type="mRNA"
             /strain="Czech II"
             /db_xref="taxon:10090"
             /clone="IMAGE:2644947"
             /tissue_type="spontaneous tumor, metastatic to mammary.
             Stem cell origin."
             /lab_host="DH10B"
             /clone_lib="NCI CGAP Lu29"
             /note="Organ: lung; Vector: pCMV-SPORT6; Site 1: SalI;
             Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
             Library constructed by Life Technologies. Investigator
             providing samples: Gilbert Smith, NIH"

ORIGIN
Alignment Scores:
Pred. No.:      3,78e-63      Length:      669
Score:          672.00      Matches:    141
Percent Similarity: 88.14%      Conservative: 15
Best Local Similarity: 79.66%      Mismatches: 15
Query Match:    61.37%      Indels:     6
DB:             10          Gaps:        2

US-10-057-510-4 (1-216) x AW231125 (1-669)

Qy 1 MetCysAsnThrAspMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 147 ATGTGCAATACCAACATGCTGTCTACCGAGGGTGTGCAAGCACCTCAGATTCCA 206
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuSer 40
Db 207 GTTTCGGACAGAGACTCTGTTGTAGACCAAAACCATTCGTTTGAAGTTGTTAAGTCC 266
Qy 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu---GlnTyr 59
Db 267 GTTGAGCGCAAAACGACACTTACACTATGAAGAGATTATATTATTATGGCCAGTAT 326
Qy 60 IleMetThrLysArgLeuTyThrAspGluLysGlnGlnHisIleValTyCysSerAsp 79
Db 327 ATTATGACTAGAGTTATATACAGAGAGCAGCAGCATTTGTATGTTCAAATGAT 386
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 387 CTCCTAGAGAGATGTTTGGAGTCCCGAGTTCTCTGTGAAGGAGCACAGAAATATAT 446
Qy 100 ThrMetIleTyArgAsnLeuValValValAsnGlnGlnGluSerSerAspSerGlyThr 119

```

```

Db 447 GCAATGATCTACAGAAATTTAGTGGCTGTAAGTCAGCAA-----GACTCTGCACA 497
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAsp-LeuValG 139
Db 498 TCCTGAGTGAGAGAGAGAGCTCAGCCTGAGGGTGGAGTGATCTGAGGATCCTTTTGA 557
Qy 139 n-GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSer 159
Db 558 AGCGCCCAACCAAGAGAACCTTCATCTCTGATTAAATTTCTAGACTGTCTACCTCAT 617
Qy 159 erArgArgAlaIleSerGluThrGluLysSerAspGluLeu 174
Db 618 CTAAGAGGAGATCCATTAGTGACAGAGAGAGACACAGATGAGCTA 664

RESULT 15
LOCUS    R80343/c
DEFINITION Y196d02.r1 Soares placenta Nb2HP Homo sapiens cDNA clone
            IMAGE:147075 5' similar to gb:M92424 MDM2 PROTEIN (HUMAN); mRNA
            sequence.
ACCESSION R80343
VERSION   R80343.1  GI:856624
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 482)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Eliston,K., Hawkins,M.,
            Holman,M., Hultman,M., Kucaba,T., Le.M., Lennon,G., Marra,M.,
            Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
            Trevisan,E., Waterston,R., Williamson,A., Wohlmann,P. and
            Wilson,R.
TITLE     The WashU-Merck EST Project
JOURNAL   Unpublished (1995)
COMMENT   Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 529
            High quality sequence stops: 371
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 529 Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 371.
            Location/Qualifiers
            1..482
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="GB:558688"
            /db_xref="taxon:9606"
            /clone="IMAGE:147075"
            /sex="Female"
            /dev_stage="placenta obtained at birth (full term)"
            /lab_host="DH10B (ampicillin resistant)"
            /notes="Organ: placenta; Vector: pT7T3D (Pharmacia) with a
            modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
            strand cDNA was primed with a Not I - oligo(dT) primer [5',
            AACTGGAAGAATTCGGCGCGGAGGATTTTCTTTTCTTTT 3'],
            double-stranded cDNA was ligated to Eco RI adaptors
            (Pharmacia), digested with Not I and cloned into the Not I
            and Eco RI sites of the modified pT7T3 vector. Library
            went through one round of normalization. Library
            constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Alignment Scores:

```

Pred. No.: 1.52e-62 Length: 482
Score: 664.50 Matches: 139
Percent Similarity: 96.53% Conservative: 0
Best Local Similarity: 96.53% Mismatches: 1
Query Match: 60.68% Indels: 4
DB: 14 Gaps: 1

US-10-057-510-4 (1-216) x R80343 (1-482)

```
Qy 1 MetCys-AsnThrAsnMetSerValProThrAspGlyAla-ValThrThrSerGlnIle- 19
Db 427 ATGTGCCAATACCAAAATGCTGTACCTACTGATGGTGCCTGTACCCCTCAGATTTC 368
Qy 20 ProLaserGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysLeuLys 39
Db 367 CCAGCTTCGGAACAAGAGACCCCTGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAG 308
Qy 40 SerValGlyAlaGlnLysAsnThrTyxThrMetLysGluValLeuPheTyxLeu---Gln 58
Db 307 TCTGTGTGTGCACAAAAGACACTTATACATATGAAGAGGTCTCTTTTATCTTGCCAG 248
Qy 59 TyrIleMetThrLysArgLeuTyxAspGluLysGlnGlnHisIleValTyxCysSerAsn 78
Db 247 TATATTATGACTAAACGATTATATGATGAGAGGACACACATATTGTATTGTTCAAT 188
Qy 79 AspLeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIle 98
Db 187 GATCTTCTAGGAGATTGTTTGGCGTGCCAAAGCTTCTCTGTGAAGAGACACAGGAAAATA 128
Qy 99 TyrThrMetIleTyxArgAsnLeuValValValAsnGlnGlnGlnSerSerAspSerGly 118
Db 127 TATACCATGATCTACAGGAACCTTGTTAGTAGTCAATCAGCAGGATCAICGACTCAGGT 68
Qy 119 ThrSerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuVal 138
Db 67 ACATCTGTGAGTGAGAACAGGTGTACCTTGAAGTGGGAGTGCATCAAAAGGACCTTGTA 8
Qy 139 GlnGlu 140
Db 7 CAGAG 2
```

Search completed: March 15, 2004, 00:19:16
Job time : 2683.17 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2004, 21:54:50 ; Search time 3245.21 Seconds
(without alignments)
2884.893 Million cell updates/sec

Title: US-10-057-510-4
Perfect score: 1095
Sequence: 1 MCNTNMSVFDGAVTTSQIP.....ALCVIREICCRSSSSSESTG 216

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgm2_1/USPTO_spool/US10057510/runat_09032004_162234_25975/app_query.fasta_1.654
-DB=GenEmbl -OPMT=fastap -SUFFIX=scd.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10057510 -CGN 1_1_2471 @runat_09032004_162234_25975 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.in.*
5: gb.ov.*
6: gb.ov.*
7: gb.pat.*
8: gb.ph.*
9: gb.pl.*
10: gb.pr.*
11: gb.ro.*
12: gb.sts.*
13: gb.sy.*
14: gb.un.*
15: gb.vi.*
16: em.ba.*
17: em.fun.*
18: em.hum.*
19: em.in.*
20: em.mu.*
21: em.om.*
22: em.or.*
23: em.ov.*
24: em.pat.*
25: em.ph.*
26: em.pl.*
27: em.ro.*
28: em.ste.*
29: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1084.5	99.0	652	6	AR207094 Sequence 4
2	1084.5	99.0	852	6	A44504 Sequence 4
3	1084.5	99.0	1476	6	A61359 Sequence 1
4	1084.5	99.0	1476	6	AX057138 Sequence 1
5	1084.5	99.0	1476	6	AX695559 Sequence
6	1084.5	99.0	1476	9	BT007258 Homo sapi
7	1084.5	99.0	1476	12	BT007935 Synthetic
8	1084.5	99.0	2372	6	AR000256 Sequence
9	1084.5	99.0	2372	6	AR009781 Sequence
10	1084.5	99.0	2372	6	AR028963 Sequence
11	1084.5	99.0	2372	6	ARI54584 Sequence
12	1084.5	99.0	2372	6	BD233925 Method an
13	1084.5	99.0	2372	6	BD233925 Sequence 1
14	1084.5	99.0	2372	6	I11727 Sequence 1
15	1084.5	99.0	2372	6	I12226 Sequence 1
16	1084.5	99.0	2372	6	I21256 Sequence 1
17	1084.5	99.0	2372	6	I25341 Sequence 2
18	1084.5	99.0	2372	6	I36472 Sequence 2
19	1084.5	99.0	2372	6	I40222 Sequence 2
20	1084.5	99.0	2372	6	I79856 Sequence 2
21	1084.5	99.0	2372	6	I86850 Sequence 2
22	1084.5	99.0	2372	6	AR212312 Sequence
23	1084.5	99.0	2372	6	AR214399 Sequence
24	1084.5	99.0	2372	6	AX329941 Sequence
25	1084.5	99.0	2372	6	AX587651 Sequence
26	1084.5	99.0	2372	6	AX695558 Sequence
27	1084.5	99.0	2372	6	BD073962 Antisense
28	1084.5	99.0	2372	6	BD138075 Antisense
29	1084.5	99.0	2372	9	HSP53ASSG
30	1035.5	99.0	2372	9	HUMSP53A
31	1027.5	93.8	2042	4	AB099709 Felis cat
32	1015.5	92.7	1469	4	AB031276 Canis fam
33	1011.5	92.4	1477	4	AF122416 Canis fam
34	1004.5	91.7	1460	4	AF100705 Canis fam
35	987	90.1	1564	4	AF322417 Canis fam
36	940.5	85.9	646	9	AF385323 Homo sapi
37	940.5	85.9	729	6	A61763 Sequence 3
38	940.5	85.9	729	6	AR264886 Sequence
39	860	78.5	732	9	HSAC30612 Homo sapi
40	842.5	76.9	1399	10	MAU10982 Mesocricetu
41	829	75.7	1391	9	AF092845 Homo sapi
42	808	73.8	681	6	A44505 Sequence 5
43	808	73.8	1302	6	A44506 Sequence 6
44	805	73.5	1470	10	MMU47934 Mus musculu
45	805	73.5	1711	10	MMU40145 Mus musculu

ALIGNMENTS

AR207094 AR207094 652 bp DNA linear PAT 20-JUN-2002
LOCUS
DEFINITION Sequence 3 from patent US 6372490.
ACCESSION AR207094
VERSION AR207094.1 GI:21505896
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1. (bases 1 to 652)
Nandabalan, K., Yang, M. and Schulz, V.
Nucleic acid encoding the MDN interacting protein
Patent: US 6372490-A 3 16-APR-2002;
JOURNAL Location/Qualifiers
1. 652
/organism="unknown"
/mol_type="unassigned DNA"
source
ORIGIN
Alignment Scores:
Pred. No.: 7,78e-99 Length: 652
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservatives: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1
US-10-057-510-4 (1-216) x AR207094 (1-652)
Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnilePro 20
Db 1 ATGTGCAATACCAACATGCTGTACTACTGATGCTGTGACACCTCACAGTTCCA 60
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysLeuSer 40
Db 61 GCTTCGGAAACAAGAGACCTGTTAGACCAAGCCATTGCTTTTGAAGTTATTAATCT 120
Qy 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu--GlnTy 59
Db 121 GTTGTGTCACAAAGACACTTATCTATGAAAGAGGTCTTTTATCTTGGCCAGAT 180
Qy 60 IleMetThrLysArgLeuTyAspGluLysGlnGlnHisIleValTyCysSerAsnAsp 79
Db 181 ATTATGACTAACGATTATATGATGAGAGCAACACATATGTTATTTGTTCAATGAT 240
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysileTy 99
Db 241 CTTCTAGGAGATTCTTTGGCGTGCACAGCTTCTCTGTGAAAGAGCACAGGAAATATAT 300
Qy 100 ThrMetIleTyArgAsnLeuValValAlaGlnGlnGlnSerSerAspSerGlyThr 119
Db 301 ACCATGATCTACAGAACTTGGTAGTATCAATCAGCAGGATCATCGGACTCAGGTACA 360
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139
Db 361 TCTGTGAGTGAGAACAGGTGTCACTTGAAGTGGGAGTGATCAAAAGACCTTGTACAA 420
Qy 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 421 GAGCTTCAGGAAGAGAAACCTTCATCTTCACATTTGGTTTCTAGACCATCTACCTCATCT 480
Qy 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyLysArgGln 179
Db 481 AGAAGGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATTATCTGGTGAACGACAA 540
Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 541 AGAAACCGCCACAAATCTGATAGTATTCCTTTCCTTTGATGAAGCCTCGCTCTGTGT 600
Qy 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
Db 601 GTAATAAGGGAGATATGTTGTGAAAGAGACAGTAGCAGTGAATCTACAGGG 651

RESULT 2
LOCUS A44504 852 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 4 from Patent WO9514233.
ACCESSION A44504
VERSION A44504.1 GI:2299322
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE
1. (bases 1 to 852)
Zentgraf H., Klein R., Frey M. and Martens R.
METHOD OF IDENTIFYING HDM-2-SPECIFIC ANTIBODIES
Patent: WO 9514233-A 4 26-MAY-1995;
JOURNAL DEUTSCHES KREBSFORSCH (DE)
COMMENT Other publication DE 4339533 950614
Other publication DE 4345249 950524.
FEATURES
1..852
Location/Qualifiers
source
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Alignment Scores:
Pred. No.: 1,09e-98 Length: 852
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservatives: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1
US-10-057-510-4 (1-216) x A44504 (1-852)
Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnilePro 20
Db 1 ATGTGCAATACCAACATGCTGTACTACTGATGCTGTGACACCTCACAGTTCCA 60
Qy 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysLeuSer 40
Db 61 GCTTCGGAAACAAGAGACCTGTTAGACCAAGCCATTGCTTTTGAAGTTATTAATCT 120
Qy 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu--GlnTy 59
Db 121 GTTGTGTCACAAAGACACTTATCTATGAAAGAGGTCTTTTATCTTGGCCAGAT 180
Qy 60 IleMetThrLysArgLeuTyAspGluLysGlnGlnHisIleValTyCysSerAsnAsp 79
Db 181 ATTATGACTAACGATTATATGATGAGAGCAACACATATGTTATTTGTTCAATGAT 240
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysileTy 99
Db 241 CTTCTAGGAGATTCTTTGGCGTGCACAGCTTCTCTGTGAAAGAGCACAGGAAATATAT 300
Qy 100 ThrMetIleTyArgAsnLeuValValAlaGlnGlnGlnSerSerAspSerGlyThr 119
Db 301 ACCATGATCTACAGAACTTGGTAGTATCAATCAGCAGGATCATCGGACTCAGGTACA 360
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139
Db 361 TCTGTGAGTGAGAACAGGTGTCACTTGAAGTGGGAGTGATCAAAAGACCTTGTACAA 420
Qy 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 421 GAGCTTCAGGAAGAGAAACCTTCATCTTCACATTTGGTTTCTAGACCATCTACCTCATCT 480
Qy 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyLysArgGln 179
Db 481 AGAAGGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATTATCTGGTGAACGACAA 540
Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 541 AGAAACCGCCACAAATCTGATAGTATTCCTTTCCTTTGATGAAGCCTCGCTCTGTGT 600

```

Qy 200 ValileArgGluileCysCysGluArgSerSerSerSerGluSerThrGly 216
|
|
|
Db 601 GTAATAAGGAGATATGTTGTGAAGAAGCAGTAGCAGTGAATCTACAGG 651

RESULT 3
A61359
LOCUS A61359 1476 bp DNA linear PAT 09-MAR-1998
DEFINITION Sequence 1 from Patent WO9709343.
ACCESSION A61359
VERSION A61359.1 GI:3715769
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Tocque,B., Dube-Poterszman,M. and Waslyk,B.
TITLE ANTAGONISTS OF THE ONCOGENIC ACTIVITY OF THE PROTEIN MDM2, AND USE
THEREOF IN THE TREATMENT OF CANCERS
JOURNAL Patient: WO 9709343-A 1 13-MAR-1997;
RHONE POULENC RORER SA (FR)
COMMENT Other publication FR 2738151 970307.
FEATURES
Location/Qualifiers
1..1476
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1..1476
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAA03588.1"
/db_xref="GI:3715770"
/db_xref="REMBL:CAA03588"
/translation="MCNTNMSVPTDGAATTSQIPASEQETLVRPKPLLLKLLKSVGAQ
KDTYTKVELFYLGQYIMTKRLYDEKQKHIVCSNDLLGLDFGVPSFVKHRYITM
IYNLVVNVQOESDSGTSVENRCHLEGGSDQDLVQLQEKPSHLSVRPSTSS
RRRAISETENSDLSGERQRKRHSDSISLSPDESIALCVIREICCRSSSSSTGT
PSPDLAGVSEHSGWLDQDSQDFSEVFEVSLSDSEYSLSEGGQLSDDEDEVY
QVTYVQAGSDTSPFEDPEISLADYWKCTSCNEMNPLPSHCNRCWALRENWLPEDK
GDKQISSEKAKLENSTQAEFGFVDPCKTIVNDSRESVCSLPLNAIEPCVIOQRPNKCI
DYQPSSTSSIIYSQEDVKEFERETQDKESVESLPLNAIEPCVIOQRPNKCI
VHGKTHLMACTCAKLLKRNKPCPCVCRQPIQMVLYTFP"
source
ORGANISM Homo sapiens
LOCUS AX057138 1476 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 14 from Patent WO0075184.
ACCESSION AX057138
VERSION AX057138.1 GI:12309959
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Zhang,H., Tsvetkov,L.M. and Kondo,T.
TITLE Modulation of protein levels using the scf complex
JOURNAL Patent: WO 0075184-A 14 14-DEC-2000;
YALE UNIVERSITY (US)
FEATURES
Location/Qualifiers
1..1476
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..1476
/notes="unnamed protein product; Human MDM2"
/codon_start=1
/protein_id="CAC22426.1"
/db_xref="GI:12309960"
/db_xref="REMBL:CAC22426"
/translation="MCNTNMSVPTDGAATTSQIPASEQETLVRPKPLLLKLLKSVGAQ
KDTYTKVELFYLGQYIMTKRLYDEKQKHIVCSNDLLGLDFGVPSFVKHRYITM
IYNLVVNVQOESDSGTSVENRCHLEGGSDQDLVQLQEKPSHLSVRPSTSS
RRRAISETENSDLSGERQRKRHSDSISLSPDESIALCVIREICCRSSSSSTGT
PSPDLAGVSEHSGWLDQDSQDFSEVFEVSLSDSEYSLSEGGQLSDDEDEVY
QVTYVQAGSDTSPFEDPEISLADYWKCTSCNEMNPLPSHCNRCWALRENWLPEDK
GDKQISSEKAKLENSTQAEFGFVDPCKTIVNDSRESVCSLPLNAIEPCVIOQRPNKCI
DYQPSSTSSIIYSQEDVKEFERETQDKESVESLPLNAIEPCVIOQRPNKCI
VHGKTHLMACTCAKLLKRNKPCPCVCRQPIQMVLYTFP"
source
ORIGIN
Alignment Scores:
Pred. No.: 2,19e-98 Length: 1476
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1

US-10-057-510-4 (1-216) x A61359 (1-1476)

Qy 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrSerGlnilePro 20
|
|
|
Db 1 ATGTGCAATACCAACATGTCGTACCTACTGATGGTCTGTAACCACTCACAGATTCCA 60
|
|
|
Qy 21 AlaserGluGlnThrLeuValArgProLysProleuLeuLeuLeuLysLeuLysSer 40
|
|
|
Db 61 GCTTCGGAACAAGAGACCCCTGGTGTAGACCAAGCCATTGCTTTGAAGTATTAAAGTCT 120
|
|
|
Qy 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu---GlnTy 59
|
|
|
Db 121 GTTGTGTCACAAAAGACATTATCTATGTAAGAGGTTCTTTTATCTTGGCCAGTAT 180
|
|
|
Qy 60 IleMetThrLysArgLeuTyAspGluLysGlnGlnHisIleValTyCysSerAsnAsp 79
|
|
|
Db 181 ATTATGACTAACGATTATATGATGAGAGACCAACATATTTCTATATTTGTTCAATGAT 240
|
|
|
Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgGlyIleTy 99
|
|
|
Db 241 CTCTTAGGAGATTGTGTTGGCGTCCCAAGCTTCTCTGTGAAGAGACAGGAAATATAT 300

```

```

Qy 100 ThrMetIleTyArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
|
|
|
Db 301 ACCATGATCTACAGGAACCTTGCTAGTAGTCAATCAGCAGGAATCATCGACTCAGGTACA 360
|
|
|
Qy 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
|
|
|
Db 361 TCTGTGAGTGAAGACAGGTGTCACCTTGAAGGTGGAGTGATCAAAAGACCTTGATCAAA 420
|
|
|
Qy 140 GluLeuGlnGluGluLysProSerSerSerHisLeuValSerArgProSerThrSerSer 159
|
|
|
Db 421 GAGCTTCAGGAAGAAACCTTCATCTTCACATTGGTTTCTAGACCATCTACCTCATCT 480
|
|
|
Qy 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
|
|
|
Db 481 AGAAGGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATTCCTGGTGAACGACAA 540
|
|
|
Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
|
|
|
Db 541 AGAAACCCCAAAATCTGATAGTATTTCCCTTTCTTTGATGAAGAGCTGGCTCTGTGT 600
|
|
|
Qy 200 ValileArgGluileCysCysGluArgSerSerSerSerGluSerThrGly 216
|
|
|
Db 601 GTAATAAGGAGATATGTTGTGAAGAAGCAGTAGCAGTGAATCTACAGG 651

RESULT 4
AX057138 1476 bp DNA linear PAT 17-JAN-2001
LOCUS AX057138
DEFINITION Sequence 14 from Patent WO0075184.
ACCESSION AX057138
VERSION AX057138.1 GI:12309959
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Zhang,H., Tsvetkov,L.M. and Kondo,T.
TITLE Modulation of protein levels using the scf complex
JOURNAL Patent: WO 0075184-A 14 14-DEC-2000;
YALE UNIVERSITY (US)
FEATURES
Location/Qualifiers
1..1476
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
1..1476
/notes="unnamed protein product; Human MDM2"
/codon_start=1
/protein_id="CAC22426.1"
/db_xref="GI:12309960"
/db_xref="REMBL:CAC22426"
/translation="MCNTNMSVPTDGAATTSQIPASEQETLVRPKPLLLKLLKSVGAQ
KDTYTKVELFYLGQYIMTKRLYDEKQKHIVCSNDLLGLDFGVPSFVKHRYITM
IYNLVVNVQOESDSGTSVENRCHLEGGSDQDLVQLQEKPSHLSVRPSTSS
RRRAISETENSDLSGERQRKRHSDSISLSPDESIALCVIREICCRSSSSSTGT
PSPDLAGVSEHSGWLDQDSQDFSEVFEVSLSDSEYSLSEGGQLSDDEDEVY
QVTYVQAGSDTSPFEDPEISLADYWKCTSCNEMNPLPSHCNRCWALRENWLPEDK
GDKQISSEKAKLENSTQAEFGFVDPCKTIVNDSRESVCSLPLNAIEPCVIOQRPNKCI
DYQPSSTSSIIYSQEDVKEFERETQDKESVESLPLNAIEPCVIOQRPNKCI
VHGKTHLMACTCAKLLKRNKPCPCVCRQPIQMVLYTFP"
source
ORIGIN
Alignment Scores:
Pred. No.: 2,19e-98 Length: 1476
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1

US-10-057-510-4 (1-216) x AX057138 (1-1476)

```

QY 1 MetCysAenThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnilePro 20
 Db 1 ATGTGCAATACCAACATGCTGTACTACTGATGCTGTGTAACACCTCACAGATTCCA 60
 QY 21 AlaserGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLeu 40
 Db 61 GCTTCGGAACAGAGACCCCTGTTAGACCAAGCCATTGCTTTGAAGTTATTAAAGTCT 120
 QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
 Db 121 GTTGTGTCACAAAAGACACATTATATGTAAGAGAGGTTCTTTTATCTTGGCCAGTAT 180
 QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
 Db 181 ATTATGACTAAACGATTATATGATGAGAAGCAACACATATTGTATATTGTTCAAAATGAT 240
 QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGlnHisIleValTyrCysSerAsnAsp 99
 Db 241 CTCTAGGAGATTGTTTGGCGTGCCAGCTTCTCTGTGAAGAGCAGCAGGAATATAT 300
 QY 100 ThrMetIleTyrArgAsnLeuValValLysGlnGlnGlnHisIleValTyrCysSerAsnAsp 119
 Db 301 ACCATGATCTACAGGAACCTTGGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 360
 QY 120 SerValSerGluAsnArgCysHisIleuGluGlySerAspGlnLysAspLeuValGln 139
 Db 361 TCTGTGAGTGAGAACAGGTGTCACTTGAAGTGGAGTGATCAAAAGGACCTTGTACAA 420
 QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
 Db 421 GAGCTTCAGGAAGAGAACCTTCACTTTCACATTGTTGTTCTAGACCATCTACCTCATCT 480
 QY 160 ArgArgArgAlaIleSerGluThrGluLysAsnSerAspGluLeuSerGlyGluArgGln 179
 Db 481 AGAAGCAGAGCAATTAAGTGAGACAGAGAAATTCAGATGAATTAATCTGGTGAACGACAA 540
 QY 180 ArgLysArgHisLysSerAspSerLysLeuSerPheAspGluSerLeuAlaLeuCys 199
 Db 541 AGAAGCAGAGCAATTAAGTGAGATGATATTCCTTCTTCTTGTGAAGCTGGCTCTGTGT 600
 QY 200 ValIleArgGluLeuCysGluArgSerSerSerSerSerSerSerSerSerSerSer 216
 Db 601 GTAATAAGGGAGATATGTTGTGAAGAGCAGTAGCAGTGAATCTACAGGG 651

RESULT 5
 AX695559
 LOCUS AX695559 1476 bp DNA linear PAT 31-MAR-2003
 DEFINITION Sequence 1186 from Patent WO03008583.
 ACCESSION AX695559
 VERSION AX695559.1 GI:29418711

KEYWORDS Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1

REFERENCE Morris D.W. and Engelhard, E.K.
 Novel compositions and methods for cancer
 Patent: WO 03008583-A 1186 30-JAN-2003;
 JOURNAL Sagres Discovery (US)

FEATURES
 Location/Qualifiers
 1..1476
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2,19e-98 Length: 1476
 Score: 1084.50 Matches: 216
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 0
 Query Match: 99.04% Indels: 1

Db: 6 Gaps: 1

US-10-057-510-4 (1-216) x AX695559 (1-1476)

QY 1 MetCysAenThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnilePro 20

Db 1 ATGTGCAATACCAACATGCTGTACTACTGATGCTGTGTAACACCTCACAGATTCCA 60

QY 21 AlaserGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeu 40

Db 61 GCTTCGGAACAGAGACCCCTGTTAGACCAAGCCATTGCTTTGAAGTTATTAAAGTCT 120

QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59

Db 121 GTTGTGTCACAAAAGACACATTATATGTAAGAGAGGTTCTTTTATCTTGGCCAGTAT 180

QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79

Db 181 ATTATGACTAAACGATTATATGATGAGAAGCAACACATATTGTATATTGTTCAAAATGAT 240

QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGlnHisIleValTyrCysSerAsnAsp 99

Db 241 CTCTAGGAGATTGTTTGGCGTGCCAGCTTCTCTGTGAAGAGCAGCAGGAATATAT 300

QY 100 ThrMetIleTyrArgAsnLeuValValLysGlnGlnGlnHisIleValTyrCysSerAsnAsp 119

Db 301 ACCATGATCTACAGGAACCTTGGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 360

QY 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139

Db 361 TCTGTGAGTGAGAACAGGTGTCACTTGAAGTGGAGTGATCAAAAGGACCTTGTACAA 420

QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159

Db 421 GAGCTTCAGGAAGAGAACCTTCACTTTCACATTGTTGTTCTAGACCATCTACCTCATCT 480

QY 160 ArgArgArgAlaIleSerGluThrGluLysAsnSerAspGluLeuSerGlyGluArgGln 179

Db 481 AGAAGCAGAGCAATTAAGTGAGACAGAGAAATTCAGATGAATTAATCTGGTGAACGACAA 540

QY 180 ArgLysArgHisLysSerAspSerLysLeuSerPheAspGluSerLeuAlaLeuCys 199

Db 541 AGAAGCAGAGCAATTAAGTGAGATGATATTCCTTCTTCTTGTGAAGCTGGCTCTGTGT 600

QY 200 ValIleArgGluLeuCysGluArgSerSerSerSerSerSerSerSerSerSerSer 216

Db 601 GTAATAAGGGAGATATGTTGTGAAGAGCAGTAGCAGTGAATCTACAGGG 651

RESULT 6

BT007258

LOCUS

DEFINITION

BT007258

ACCENSION

BT007258.1

KEYWORDS

FLI CDNA.

ORGANISM

Homo sapiens

Homo sapiens

REFERENCE

1 (bases 1 to 1476)

Cloning of human full-length CDSs in BD Creator(TM) System Donor

vector

Unpublished

2 (bases 1 to 1476)

REFERENCE

Kalnina, N., Chen, X., Rolfs, A., Halleck, A., Hines, L., Eisenstein, S.,

Koundinya, M., Raphael, J., Moreira, D., Kelley, T., LaBaer, J., Lin, Y.,

Phelan, M. and Farmer, A.

Direct Submission

Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow

JOURNAL

KDTYTMKEVLYFGYIMTKRLYDEKQHIYVCSNDLGLDFGVPSFVKEHRKIYTM
IYRLNVVNOQSSDSGTSENRKHLGGSDQDLVQELQEBKPSHLSVSRPSTSS
RRRAISEVNSDELSGRQRKRKSDSISLDFESIALCVIRICEGSSSESTGT
PSNPDLGAVSEHSDMLDQSDVSQDFSEVESLDSQYSLSEGEQLSDRDEYV
QVTVYQAGSESTDEEDPEISLDYKCTSCNEMWPLPSHCNRCWALRENWLPEDK
GKDGEISEKALNISTQABEGDFVDCCKTIIVDSRESCEENDDKTIQASQSESE
DYQSTSSSIYSSQEDVBEFEETODKEESVESLPLNLAIBPCVICQGRPNKGI
VHGKTHLMACFTCAKLLKRNKPCVCRQPIQIMVLTPEL"

ORIGIN

Alignment Scores:
Pred. No.: 2,198-98 Length: 1476
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 12 Gaps: 1

US-10-057-510-4 (1-216) x BT007935 (1-1476)

QY 1 MetCysAenThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 1 ATGTGCAATACCAACATCTGTACTACTGATGCTGTAAACACCTCACAGATTCCA 60
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLeu 40
Db 61 GCTTCGGAACAAGAGACCCCTGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 121 GTTGTGCACAAAAGACACATATATGATGAAGAGGTTCTTTTATCTTGGCCAGTAT 180
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnHisIleValTyrCysSerAsnAsp 79
Db 181 ATTATGACTAACCGATTATATGATGAGAGCAACAACATATTTGTTTCAATGAT 240
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 241 CTCTTAGAGATTGTTTGGCGTCCAGCTTCTCTGTGAAGAGCAGCAAGAAATATAT 300
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnLysSerAspSerGlyThr 119
Db 301 ACCATGATCTACAGAACTTGATAGTATCAATCAGCAGGAATCATCGGACTCAGGTACA 360
QY 120 SerValSerGluAenArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 361 TCTGTGAGTGAGAACAGGTGTCACCTTGAAGTGGAGTGATCAAAAGGACCTTGTACAA 420
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 421 GAGCTTCAGGAAGAGAAACCTTCATCTTCACATTTGTTTCTAGACCATCTACCTCATCT 480
QY 160 ArgArgArgAlaIleSerGluThrGluGluAenSerAspGluLeuSerGlyGluArgGln 179
Db 481 AGAAGGAGAGCAATTAGTGAGACAGAAAGAAATTCAGATGAATTTATCTGTGACGACAA 540
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 541 AGAAGACCCCAAAATCTGATAGTATTTCCCTTTTCCCTTTGATGAAGACCTTGTCTGTGT 600
QY 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
Db 601 GTAATAAGGAGATATGTTGTAAGAGCAGTACGATGAATCTACAGGG 651

RESULT 8

AR000256 LOCUS 2372 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 2 from patent US 5736338.
ACCESSION AR000256
VERSION AR000256.1 GI:3962787
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.

Unclassified.
1 (bases 1 to 2372)
Burrell, M., Hill, D.E., Kinzler, K.W. and Vogelstein, B.
Method of diagnosing Neoplastic disease by detecting increased
expression of human MDM2 protein
Patent: US 5736338-A 2 07-APR-1998;
Location/Qualifiers
1..2372
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3,986-98 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1

US-10-057-510-4 (1-216) x AR000256 (1-2372)

QY 1 MetCysAenThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATCTGTACTACTGATGCTGTAAACACCTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLysSer 40
Db 372 GCTTCGGAACAAGAGACCCCTGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 432 GTTGTGCACAAAAGACACCTTATATGATGAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnHisIleValTyrCysSerAsnAsp 79
Db 492 ATTATGACTAACCGATTATATGATGAGAGCAACAACATATTTGTTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 552 CTCTTAGAGATTGTTTGGCGTCCAGCTTCTCTGTGAAGAGCAGCAAGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnLysSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGAACTTGATAGTATCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAenArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTGAGTGAGAACAGGTGTCACCTTGAAGTGGAGTGATCAAAAGGACCTTGTACAA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GAGCTTCAGGAAGAGAAACCTTCATCTTCACATTTGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluGluAenSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATTAGTGAGACAGAAAGAAATTCAGATGAATTTATCTGTGACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAGACCCCAAAATCTGATAGTATTTCCCTTTTCCCTTTGATGAAGACCTTGTCTGTGT 911
QY 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
Db 912 GTAATAAGGAGATATGTTGTAAGAGCAGTACGATGAATCTACAGGG 962

RESULT 9

AR0009781 LOCUS 2372 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 2 from patent US 5756455.
ACCESSION AR0009781
VERSION AR0009781.1 GI:3968586
KEYWORDS

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2372)
AUTHORS Kinzler, K.W. and Vogelstein, B.
TITLE Amplification of human MDM2 gene in human tumors
JOURNAL Patent: US 5756455-A 2 26-MAY-1998;
FEATURES Location/Qualifiers
1. 2372
source /organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3.98e-98 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1

US-10-057-510-4 (1-216) x AR009781 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACCTACTGATGGTGTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLysSer 40
DB 372 GCCTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 432 GTTGGTGCACAAAAGACACTTACTATGAAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 492 ATTATGACTAAACGATTATATGATGAGAGCAACACATATTGTATTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTTCTAGGAGATTGTTTGGCGTGCACAGCTTCTGTGAAAGAGCACAGGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnLysSerSerAspSerGlyThr 119
DB 612 ACCATGATCTACAGAACTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyCysSerAspGlnLysAspLeuValGln 139
DB 672 TCTGTGAGTGAGAACAGGTGTCACCTTGAAGTGGGAGTGATCAAAAGGACCTTGACAA 731
QY 140 GluLeuGlnGluLysProSerSerSerHisLeuValSerArgProSerThrSerSer 159
DB 732 GAGCTTCAGAGAGAAACCTTCATCTTCATTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluAsnSerAspGluLeuSerGlyGluArgGln 179
DB 792 AGAAGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATTATCTGGTGAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 852 AGAAACGCCACAAATCTGATAGTATTTCCTTTTCCCTTTGATGAAAGCCTGGCTGTGT 911
QY 200 ValIleArgGluIleCysGluArgSerSerSerGluSerThrGly 216
DB 912 GTATTAAGGAGATATGTTGTGAAGAGCAGTAGCAGTGAATCTACAGGG 962

RESULT 10
AR028963
LOCUS AR028963 2372 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 2 from patent US 5858976.
ACCESSION AR028963
VERSION AR028963.1 GI:5940936

KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2372)
AUTHORS Burrell, M., Hill, D.E., Kinzler, K.W. and Vogelstein, B.
TITLE Methods for inhibiting interaction of human MDM2 and p53
JOURNAL Patent: US 5858976-A 2 12-JAN-1999;
FEATURES Location/Qualifiers
1. 2372
source /organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
Pred. No.: 3.98e-98 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1

US-10-057-510-4 (1-216) x AR028963 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACCTACTGATGGTGTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysSer 40
DB 372 GCCTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 432 GTTGGTGCACAAAAGACACTTACTATGAAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 492 ATTATGACTAAACGATTATATGATGAGAGCAACACATATTGTATTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTTCTAGGAGATTGTTTGGCGTGCACAGCTTCTGTGAAAGAGCACAGGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnLysSerSerAspSerGlyThr 119
DB 612 ACCATGATCTACAGAACTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyCysSerAspGlnLysAspLeuValGln 139
DB 672 TCTGTGAGTGAGAACAGGTGTCACCTTGAAGTGGGAGTGATCAAAAGGACCTTGACAA 731
QY 140 GluLeuGlnGluLysProSerSerSerHisLeuValSerArgProSerThrSerSer 159
DB 732 GAGCTTCAGAGAGAAACCTTCATCTTCATTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluAsnSerAspGluLeuSerGlyGluArgGln 179
DB 792 AGAAGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATTATCTGGTGAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 852 AGAAACGCCACAAATCTGATAGTATTTCCTTTTCCCTTTGATGAAAGCCTGGCTGTGT 911
QY 200 ValIleArgGluIleCysGluArgSerSerSerGluSerThrGly 216
DB 912 GTATTAAGGAGATATGTTGTGAAGAGCAGTAGCAGTGAATCTACAGGG 962

RESULT 11
AR154584
LOCUS AR154584 2372 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 1 from patent US 6238921.
ACCESSION AR154584

```
VERSION AR154584.1 GI:15122637
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2372)
AUTHORS Miraglia, L.J., Nero, P., Graham, M.J. and Monia, B.P.
TITLE Antisense oligonucleotide modulation of human mdm2 expression
JOURNAL Patent: US 6238921-A 1 29-MAY-2001;
FEATURES
    source
        1..2372
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 3,98e-98 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservatives: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1
US-10-057-510-4 (1-216) x AR154584 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATCTGTACTACTGCTGTGTAACCACTCCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysSer 40
Db 372 GCTTCGGACACAGACACCTGTAGACCAAGCATTGCTTTGAAGTTATTAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 432 GTTGTGTCACAAAAGACACTTACTATGAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
Db 492 ATTATGACTTAACGATTATATGATGAGAACCAACATATGTTATTTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 552 CTCTTAGGAGATTGTTGGCGTCCAAAGCTTCTCTGTGAAGAGCAGACAGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAlaGlnGlnGluSerSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGAACTTGGTAGTAGTCAATACAGAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTGAGTGAGAACAGAGTGTCACCTTGAGGTGGAGTGATCAAAAGACCTTGTACAA 731
QY 140 GluLeuGlnGluLysProSerSerSerHisIleValSerArgProSerThrSerSer 159
Db 732 GAGCTTCAGGAGAGAAACCTTCATCTTCAATTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluGlnAsnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATTTCTGTGAGACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAACGCCCAAAATCTGATAGTATTTCCCTTTGATGAAGCCTGGCTCTGTGT 911
QY 200 VallIleArgGluIleCysGluArgSerSerSerSerGluSerThrGly 216
Db 912 GTAATAGGAGCATATGTTGTGAAGAGACAGTAGTAGTGAATCTACAGGG 962
RESULT 12
LOCUS BD233925 2372 bp DNA linear PAT 17-JUL-2003
DEFINITION Method and reagent of enhancing growth capability and preventing
```

```
replicative senescence.
BD233925
VERSION 1 GI:33043695
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2372)
AUTHORS Hannon, G.J. and Beach, D.H.
TITLE Method and reagent of enhancing growth capability and preventing
JOURNAL Patent: JP 2002530436-A 2 17-SEP-2002;
FEATURES
    source
        1..2372
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 3,98e-98 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservatives: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1
US-10-057-510-4 (1-216) x BD233925 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATCTGTACTACTGCTGTGTAACCACTCCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysSer 40
Db 372 GCTTCGGACACAGACACCTGTAGACCAAGCATTGCTTTGAAGTTATTAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 432 GTTGTGTCACAAAAGACACTTACTATGAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
Db 492 ATTATGACTTAACGATTATATGATGAGAACCAACATATGTTATTTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 552 CTCTTAGGAGATTGTTGGCGTCCAAAGCTTCTCTGTGAAGAGCAGACAGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAlaGlnGlnGluSerSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGAACTTGGTAGTAGTCAATACAGAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTGAGTGAGAACAGAGTGTCACCTTGAGGTGGAGTGATCAAAAGACCTTGTACAA 731
```


QY 140 GluLeuGlnGluGlyProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GAGCTTCAGGAAGAGAACCTTCATCTTCACATTTGGTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATTTAGTGACAGAGAAATTCAGATGAATTCATCTGGTGAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAAAGCCACAAATCTGATAGTATTTCCCTTTCTTGTGAAAGCCTGGCTCTGTGT 911
QY 200 ValIleArgGluIleCysGluArgSerSerSerGluSerThrGly 216
Db 912 GTAATAAGGGAGATATGTTGTGAAGAAGCAGTAGCAGTGAATCTACAGGG 962

RESULT 13
111727
LOCUS 111727 2372 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 1 from Patent US 5411860.
ACCESSION 111727
VERSION 111727.1 GI:909479
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2372)
AUTHORS Vogelstein,B. and Kinzler,K.W.
TITLE Amplification of human MDM2 gene in human tumors
JOURNAL Patent: US 5411860-A 1 02-MAY-1995;
FEATURES Location/Qualifiers
1..2372
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3.98e-98 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1

US-10-057-510-4 (1-216) x 111727 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATGCTGTACCTACTGTGCTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysSer 40
Db 372 GCTTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATGCTTTGAAGTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 432 GTTGGTCGCAAAAAGACACTTATCTATGAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnHisIleValTyrCysSerAsnAsp 79
Db 492 ATTATGACTAAACGATTATATGATGAGAAGCAACAACATATTGTATTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 552 CTCTTAGAGATTGTTGGCTGCCAGCTTCTCTGTGAAGAGCACAGAAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGlnSerSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGGAACCTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139

Db 672 TCTGTGAGTGAGAACAGGTGTCACTTGAAGGTGGAGTGATCAAAAGACCTGTACAA 731
QY 140 GluLeuGlnGluGlyProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GAGCTTCAGGAAGAGAACCTTCATCTTCACATTTGGTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATTTAGTGACAGAGAAATTCAGATGAATTCATCTGGTGAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAAAGCCACAAATCTGATAGTATTTCCCTTTCTTGTGAAAGCCTGGCTCTGTGT 911
QY 200 ValIleArgGluIleCysGluArgSerSerSerGluSerThrGly 216
Db 912 GTAATAAGGGAGATATGTTGTGAAGAAGCAGTAGCAGTGAATCTACAGGG 962

RESULT 14
112226
LOCUS 112226 2372 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 1 from patent US 5420263.
ACCESSION 112226
VERSION 112226.1 GI:909724
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2372)
AUTHORS Burrell,M., Hill,D.E., Kinzler,K.W. and Vogelstein,B.
TITLE Amplification of human MDM2 Gene in human tumors
JOURNAL Patent: US 5420263-A 1 30-MAY-1995;
FEATURES Location/Qualifiers
1..2372
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3.98e-98 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 6 Gaps: 1

US-10-057-510-4 (1-216) x 112226 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATGCTGTACCTACTGTGCTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLysSer 40
Db 372 GCTTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATGCTTTGAAGTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 432 GTTGGTCGCAAAAAGACACTTATCTATGAAGAGGTTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnHisIleValTyrCysSerAsnAsp 79
Db 492 ATTATGACTAAACGATTATATGATGAGAAGCAACAACATATTGTATTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 552 CTCTTAGAGATTGTTGGCTGCCAGCTTCTCTGTGAAGAGCACAGAAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGlnSerSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGGAACCTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:45:48 ; Search time 52.8828 seconds
(without alignments)
1154.067 Million cell updates/sec

Title: US-10-057-510-4
Perfect score: 216
Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIREICCRSSSSSESTG 216

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	216	100.0	216	3 AAB08846	A human M
2	159	73.6	227	2 AAR75495	A human M
3	159	73.6	227	2 AAR75398	A human M
4	159	73.6	284	2 AAR75494	A human M
5	159	73.6	284	2 AAR75397	A human M
6	159	73.6	434	2 AAR75496	A human M
7	159	73.6	434	2 AAR75399	A human M
8	159	73.6	491	2 AAR42175	A human MDM
9	159	73.6	491	2 AAR76696	A human MDM
10	159	73.6	491	2 AAW07887	A human MDM
11	159	73.6	491	2 AAW15463	A human MDM
12	159	73.6	491	2 AAW13380	A human MDM
13	159	73.6	491	2 AAW13600	A human MDM
14	159	73.6	491	2 AAW48241	A human MDM
15	159	73.6	491	2 AAW57241	A human MDM
16	159	73.6	491	2 AAW42879	A human MDM
17	159	73.6	491	2 AAW42971	A human MDM
18	159	73.6	491	2 AAW94304	A human MDM
19	159	73.6	491	3 AAY96567	A human MDM
20	159	73.6	491	4 AAB48284	A human MDM
21	159	73.6	491	5 AAE22654	A human MDM
22	159	73.6	491	5 AAE22698	A human MDM
23	159	73.6	491	5 AAE25913	A human MDM
24	159	73.6	491	5 AAO15376	A human MDM
25	159	73.6	491	7 ADD21815	A human MDM

26	159	73.6	491	7 ADE61562	Human Pro
27	159	73.6	1171	4 AAU32421	Novel hum
28	57	26.4	205	4 ABG29958	Novel hum
29	54	25.0	361	4 ABG29956	Novel hum
30	42	19.4	59	7 ADC22294	Protein B
31	29	13.4	139	4 AAU30011	Novel hum
32	29	13.4	254	4 ABG29954	Novel hum
33	29	13.4	489	2 AAR42176	Murine MD
34	29	13.4	489	2 AAR76697	Mouse MDM
35	29	13.4	489	2 AAW07888	Murine MD
36	29	13.4	489	2 AAW15464	Murine MD
37	29	13.4	489	2 AAW48242	Mouse MDM
38	29	13.4	489	2 AAW57246	Mouse MDM
39	29	13.4	489	2 AAW42997	Amino aci
40	29	13.4	489	2 AAW42972	Amino aci
41	29	13.4	489	2 AAW94305	Mouse MDM
42	29	13.4	489	5 AAE25914	Mouse dou
43	29	13.4	489	5 ABE57099	Mouse iSC
44	29	13.4	489	5 AAO15375	Mouse Dm2
45	29	13.4	489	7 ADD21816	Mouse mdm

ALIGNMENTS

RESULT 1
AAB08846
ID AAB08846 standard; protein; 216 AA.
XX
AC AAB08846;
XX
DT 02-JAN-2001 (first entry)
XX
DE A human MDMIP-binding MDM2 polypeptide fragment.
XX
KW Human; MDM2 interacting polypeptide; MDMIP; MDM2; cell cycle progression;
KW cell differentiation; cancer; sarcoma; glioma; squamous cell carcinoma;
KW breast cancer; astrocytoma; leukemia; lymphoma; tumorogenesis;
KW gene therapy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 58 /note= "Gln encoded by GGCCAG"
FT
XX
PN WO200050590-A1.
XX
PD 31-AUG-2000.
XX
PF 23-FEB-2000; 2000WO-US004582.
XX
PR 23-FEB-1999; 99US-0121192P.
PR 03-MAR-1999; 99US-0122643P.
PR 22-FEB-2000; 2000US-00510252.
XX
PA (CUPA-) CURAGEN CORP.
XX
PI Nandabalan K, Yang M, Schulz VP;
XX
DR WPI; 2000-558398/51.
DR N-PSDB; AAA75042.
XX
PT Novel MDM2 interacting protein useful for treating or preventing
PT disorders involving aberrant levels of MDM2 and/or MDM-interacting
PT proteins, comprises a specific amino acid sequence.
PS Claim 11; Fig 2B; 78pp; English.
XX
CC The present sequence represents a fragment of a human MDM2 polypeptide,
CC which binds to a human MDM2 interacting polypeptide (MDMIP). The protein
CC fragment was used as bait in a yeast two hybrid system to identify MDMIP.
CC The MDMIP polypeptide is useful for detecting and removing MDM2

CC polypeptides in a biological sample by forming MDM2-MDMIP complexes.
 CC MDMIP and MDM2 are useful to identify compounds or other agents which
 CC modulate the activity of MDM2 and/or MDMIP-mediated processes. Agents
 CC that modulate the function of MDMIP/MDM2 complexes are useful for
 CC treating and preventing a disease or disorder involving aberrant levels
 CC of MDM2 or MDMIP. MDMIP is also useful for treating diseases caused by
 CC aberrant levels of expression of MDM2 genes, such as disorders of cell
 CC cycle progression, cell differentiation, and transcriptional control,
 CC including cancers such as human sarcoma, glioma, squamous cell carcinoma,
 CC breast cancer, astrocytoma, leukemia and lymphoma, and tumorigenesis.
 CC MDMIP and MDM2 nucleic acids are useful in gene therapy
 XX
 XX Sequence 216 AA;

Query Match 100.0%; Score 216; DB 3; Length 216;
 Best Local Similarity 100.0%; Pred. No. 6e-203;
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCNTNMSVPTDCAVTTTQIPASEQETLVPRKPLLLKLLKSVGAOKDYTKVEVLFYIYI 60
 DB 1 MCNTNMSVPTDCAVTTTQIPASEQETLVPRKPLLLKLLKSVGAOKDYTKVEVLFYIYI 60
 QY 61 MTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLLVNNQESSDSGTS 120
 DB 61 MTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLLVNNQESSDSGTS 120
 QY 121 VSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENSDELGSRQR 180
 DB 121 VSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENSDELGSRQR 180
 QY 181 KRKHSDSISLSPDESIALCVIREICCCRSSSESTG 216
 DB 181 KRKHSDSISLSPDESIALCVIREICCCRSSSESTG 216

RESULT 2
 AAR75495
 ID AAR75495 standard; protein; 227 AA.
 XX
 AC AAR75495;
 XX
 DT 02-FEB-1996 (first entry)
 XX
 DE Human double minute 2 (hdm-2) antibody-binding region fragment 2.
 XX
 KW Human double minute gene 2; hdm-2; antibody binding region; antigen;
 KW cancer; sarcoma; rhabdomyosarcoma; diagnosis; immunoassay.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..227
 FT /note= "amino acids 58-284 of hdm-2 gene product"
 XX
 PN DE4339533-Al.
 XX
 PD 14-JUN-1995.
 XX
 PF 19-NOV-1993; 93DE-04339533.
 XX
 PR 19-NOV-1993; 93DE-04339533.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 PI Zentgraf H, Klein R, Frey M, Martens R;
 XX
 DR WPI; 1995-216248/29.
 DR N-PSDB; AAQ92516.
 XX

Detection of human double minute gene 2 (hdm-2) antibodies - by
 PT incubation with new hdm-2 or antibody-binding hdm-2 fragments; useful in
 PT the detection of specific cancers.
 XX

PS Claim 11; Fig 1; 12pp; German.
 XX
 CC Fragments contg. amino acids 1-284, 58-284 and 58-491 of the hdm-2 (human
 CC double minute 2) gene product are claimed. The overlapping protein
 CC fragments contain binding regions for hdm-2- specific antibodies and are
 CC useful for identifying such antibodies in a claimed immunoassay. The
 CC presence of anti-hdm-2 antibodies is diagnostic of certain forms of
 CC cancer, e.g. rhabdomyosarcoma
 XX
 XX Sequence 227 AA;

Query Match 73.6%; Score 159; DB 2; Length 227;
 Best Local Similarity 100.0%; Pred. No. 4.1e-147;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLLVNNQESSSPS 117
 DB 2 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLLVNNQESSSPS 61
 QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENSDELGSGE 177
 DB 62 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENSDELGSGE 121
 QY 178 RQKRRKHSISLSPDESIALCVIREICCCRSSSESTG 216
 DB 122 RQKRRKHSISLSPDESIALCVIREICCCRSSSESTG 160

RESULT 3
 AAR75398
 ID AAR75398 standard; protein; 227 AA.
 XX
 AC AAR75398;
 XX
 DT 25-MAR-2003 (revised)
 DT 25-JAN-1996 (first entry)
 XX
 DE Human double minute 2 (hdm-2) antibody-binding region fragment 2.
 XX
 KW Human double minute gene 2; hdm-2; antibody binding region; antigen;
 KW cancer; sarcoma; rhabdomyosarcoma; diagnosis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..227
 FT /note= "amino acids 58-284 of hdm-2 gene product"
 XX
 PN DE4345249-Al.
 XX
 PD 24-MAY-1995.
 XX
 PF 19-NOV-1993; 93DE-04345249.
 XX
 PR 19-NOV-1993; 93DE-04339533.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 PI Zentgraf H, Klein R, Frey M, Martens R;
 XX
 DR WPI; 1995-195167/26.
 DR N-PSDB; AAQ87262.
 XX
 PT New hdm-2 fragments contg. antibody binding region - used to detect
 PT specific antibodies for diagnosis of cancers, also new DNA sequences
 PT encoding them.
 XX
 XX Claim 2; Fig 1; 11pp; German.

CC Fragments contg. amino acids 1-284, 58-284 and 58-491 of the hdm-2 (human
 CC double minute 2) gene product are claimed. The overlapping protein
 CC fragments contain binding regions for hdm-2- specific antibodies and are
 CC useful for identifying such antibodies. The presence of anti-hdm-2

CC antibodies is diagnostic of certain forms of cancer, e.g.
 CC rhabdomyosarcoma. (Updated on 25-MAR-2003 to correct Pf field.)
 XX
 SQ Sequence 227 AA;

Query Match 73.6%; Score 159; DB 2; Length 227;
 Best Local Similarity 100.0%; Pred. No. 4.1e-147;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVVNQSSDS 117
 DB 2 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVVNQSSDS 61
 QY 118 GTSVSENCHLEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISETEENSDELGE 177
 DB 62 GTSVSENCHLEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISETEENSDELGE 121
 QY 178 RQRKHKSISLSFDESALCVIREICCRSSSSESTG 216
 DB 122 RQRKHKSISLSFDESALCVIREICCRSSSSESTG 160

RESULT 4
 AAR75494
 ID AAR75494 standard; protein; 284 AA.
 XX
 AC AAR75494;
 DT 02-FEB-1996 (first entry)
 XX Human double minute 2 (hdm-2) antibody-binding region fragment 1.
 XX Human double minute gene 2; hdm-2; antibody binding region; antigen;
 KW cancer; sarcoma; rhabdomyosarcoma; diagnosis; immunoassay.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Region 1. .284
 FT /note= "amino acids 1-284 of hdm-2 gene product"

DE4339533-AL.
 PN
 PD 14-JUN-1995.
 XX
 PF 19-NOV-1993; 93DE-04339533.
 XX
 PR 19-NOV-1993; 93DE-04339533.
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX Zentgraf H, Klein R, Frey M, Martens R;
 DR WPI; 1995-216248/29.
 DR N-PSDB; AAQ92515.
 XX
 PT Detection of human double minute gene 2 (hdm-2) antibodies - by
 PT incubation with new hdm-2 or antibody-binding hdm-2 fragments; useful in
 PT the detection of specific cancers.
 XX
 XX Claim 11; Fig 1; 12pp; German.
 PS
 CC Fragments contg. amino acids 1-284, 58-284 and 58-491 of the hdm-2 (human
 CC double minute 2) gene product are claimed. The overlapping protein
 CC fragments contain binding regions for hdm-2- specific antibodies and are
 CC useful for identifying such antibodies in a claimed immunoassay method.
 CC The presence of anti-hdm-2 antibodies is diagnostic of certain forms of
 CC cancer, e.g. rhabdomyosarcoma
 XX
 SQ Sequence 284 AA;

Query Match 73.6%; Score 159; DB 2; Length 284;
 Best Local Similarity 100.0%; Pred. No. 5e-147;

Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVVNQSSDS 117
 DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVVNQSSDS 118
 QY 118 GTSVSENCHLEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISETEENSDELGE 177
 DB 119 GTSVSENCHLEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISETEENSDELGE 178
 QY 178 RQRKHKSISLSFDESALCVIREICCRSSSSESTG 216
 DB 179 RQRKHKSISLSFDESALCVIREICCRSSSSESTG 217

RESULT 5
 AAR75397
 ID AAR75397 standard; protein; 284 AA.
 XX
 AC AAR75397;
 DT 25-MAR-2003 (revised)
 DT 25-JAN-1996 (first entry)
 XX Human double minute 2 (hdm-2) antibody-binding region fragment 1.
 DE Human double minute gene 2; hdm-2; antibody binding region; antigen;
 XX cancer; sarcoma; rhabdomyosarcoma; diagnosis.
 KW Homo sapiens.
 XX Key Location/Qualifiers
 FT Region 1. .284
 FT /note= "amino acids 1-284 of hdm-2 gene product"

DE4345249-AL.
 PN
 PD 24-MAY-1995.
 XX
 PF 19-NOV-1993; 93DE-04345249.
 XX
 PR 19-NOV-1993; 93DE-04339533.
 XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX Zentgraf H, Klein R, Frey M, Martens R;
 PI WPI; 1995-195167/26.
 DR N-PSDB; AAQ87261.
 XX
 PT New hdm-2 fragments contg. antibody binding region - used to detect
 PT specific antibodies for diagnosis of cancers, also new DNA sequences
 PT encoding them.
 XX
 XX Claim 2; Fig 1; 11pp; German.
 PS
 CC Fragments contg. amino acids 1-284, 58-284 and 58-491 of the hdm-2 (human
 CC double minute 2) gene product are claimed. The overlapping protein
 CC fragments contain binding regions for hdm-2- specific antibodies and are
 CC useful for identifying such antibodies. The presence of anti-hdm-2
 CC antibodies is diagnostic of certain forms of cancer, e.g.
 CC rhabdomyosarcoma. (Updated on 25-MAR-2003 to correct Pf field.)
 XX
 SQ Sequence 284 AA;

Query Match 73.6%; Score 159; DB 2; Length 284;
 Best Local Similarity 100.0%; Pred. No. 5e-147;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVVNQSSDS 117
 DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVVNQSSDS 118

QY 118 GTSVSENCHLEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
 Db |||||
 QY 178 RQRKHKSDSLSLSPDESIALCVIREICCRSSSSSESTG 216
 Db |||||
 QY 179 RQRKHKSDSLSLSPDESIALCVIREICCRSSSSSESTG 217
 Db |||||

RESULT 6
 AAR75496
 ID AAR75496 standard; protein; 434 AA.
 XX
 AC
 XX
 DT 02-FEB-1996 (first entry)
 XX
 DE Human double minute 2 (hdm-2) antibody-binding region fragment 3.
 XX
 KW Human double minute gene 2; hdm-2; antibody binding region; antigen;
 KW cancer; sarcoma; rhabdomyosarcoma; diagnosis; immunoassay.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Region
 FT 1..434
 FT /note= "amino acids 58-491 of hdm-2 gene product"
 XX
 PN DE4339533-A1.
 XX
 PD 14-JUN-1995.
 XX
 PF 19-NOV-1993; 93DE-04339533.
 XX
 PR 19-NOV-1993; 93DE-04339533.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 PI Zentgraf H, Klein R, Frey M, Martens R;
 XX WPI; 1995-216248/29.
 DR N-PSDB; AAQ92517.
 XX
 PT Detection of human double minute gene 2 (hdm-2) antibodies - by
 PT incubation with new hdm-2 or antibody-binding hdm-2 fragments; useful in
 PT the detection of specific cancers.
 XX
 PS Claim 11; Fig 1; 12pp; German.
 XX
 CC Fragments contg. amino acids 1-284, 58-284 and 58-491 of the hdm-2 (human
 CC double minute 2) gene product are claimed. The overlapping protein
 CC fragments contain binding regions for hdm-2 specific antibodies and are
 CC useful for identifying such antibodies in a claimed immunoassay method.
 CC The presence of anti-hdm-2 antibodies is diagnostic of certain forms of
 CC cancer, e.g. rhabdomyosarcoma
 XX
 SQ Sequence 434 AA;
 Query Match 73.6%; Score 159; DB 2; Length 434;
 Best Local Similarity 100.0%; Pred. No. 7.4e-147;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKHEHKIYTIYRNLVVYNQOESSDS 117
 Db 2 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKHEHKIYTIYRNLVVYNQOESSDS 61
 QY 118 GTSVSENCHLEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
 Db |||||
 QY 62 GTSVSENCHLEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 121
 Db |||||
 QY 178 RQRKHKSDSLSLSPDESIALCVIREICCRSSSSSESTG 216
 Db |||||
 QY 179 RQRKHKSDSLSLSPDESIALCVIREICCRSSSSSESTG 217
 Db |||||

RESULT 7
 AAR75399
 ID AAR75399 standard; protein; 434 AA.
 XX
 AC
 XX
 DT 25-MAR-2003 (revised)
 DT 25-JAN-1996 (first entry)
 XX
 DE Human double minute 2 (hdm-2) antibody-binding region fragment 3.
 XX
 KW Human double minute gene 2; hdm-2; antibody binding region; antigen;
 KW cancer; sarcoma; rhabdomyosarcoma; diagnosis.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Region
 FT 1..434
 FT /note= "amino acids 58-491 of hdm-2 gene product"
 XX
 PN DE4345249-A1.
 XX
 PD 24-MAY-1995.
 XX
 PF 19-NOV-1993; 93DE-04345249.
 XX
 PR 19-NOV-1993; 93DE-04339533.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 PI Zentgraf H, Klein R, Frey M, Martens R;
 XX WPI; 1995-195167/26.
 DR N-PSDB; AAQ87263.
 XX
 PT New hdm-2 fragments contg. antibody binding region - used to detect
 PT specific antibodies for diagnosis of cancers, also new DNA sequences
 PT encoding them.
 XX
 PS Claim 2; Fig 1; 11pp; German.
 XX
 CC Fragments contg. amino acids 1-284, 58-284 and 58-491 of the hdm-2 (human
 CC double minute 2) gene product are claimed. The overlapping protein
 CC fragments contain binding regions for hdm-2 specific antibodies and are
 CC useful for identifying such antibodies. The presence of anti-hdm-2
 CC antibodies is diagnostic of certain forms of cancer, e.g.
 CC rhabdomyosarcoma. (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 434 AA;
 Query Match 73.6%; Score 159; DB 2; Length 434;
 Best Local Similarity 100.0%; Pred. No. 7.4e-147;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKHEHKIYTIYRNLVVYNQOESSDS 117
 Db 2 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPFSVKHEHKIYTIYRNLVVYNQOESSDS 61
 QY 118 GTSVSENCHLEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
 Db |||||
 QY 62 GTSVSENCHLEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 121
 Db |||||
 QY 178 RQRKHKSDSLSLSPDESIALCVIREICCRSSSSSESTG 216
 Db |||||
 QY 179 RQRKHKSDSLSLSPDESIALCVIREICCRSSSSSESTG 217
 Db |||||

RESULT 8
 AAR42175
 ID AAR42175 standard; protein; 491 AA.
 XX

AC AAR42175;
 XX 25-MAR-2003 (revised)
 DT 05-MAY-1994 (first entry)
 XX Human MDM2.
 DE
 XX p53 gene; tumour suppressor gene; regulation; cellular proliferation;
 KW cellular transformation; carcinoma; human; tumour; MDM2; inhibition;
 KW gene amplification.
 XX
 XX Homo sapiens.
 OS
 XX W09320238-A2.
 PN
 XX 14-OCT-1993.
 PD
 XX 07-APR-1993; 93WO-US003199.
 PF
 XX 07-APR-1992; 92US-00867840.
 PR 23-JUN-1992; 92US-00903103.
 PR
 XX (UYJO) UNIV JOHNS HOPKINS.
 PA
 XX Burrell M, Hill DE, Kinzler KW, Vogelstein B;
 PI WPI; 1993-336944/42.
 XX N-PSDB; AAR49891.
 DR
 XX Diagnosing neoplasia from amplification of MDM2 gene - or elevated gene
 PT expression, also new DNA, MDM2 protein, antibodies and treatment of
 PT sarcoma by inhibiting MDM2 expression.
 XX
 XX Claim 19; Fig 1; 75pp; English.
 PS
 XX This sequence is encoded by the MDM2 gene. Amplification of the MDM2 gene
 CC is diagnostic of neoplasia or the potential for neoplasia. The protein
 CC encoded by this gene interacts with the product of the p53 gene. p53 is a
 CC tumour suppressor gene and encodes a protein which appears to be a member
 CC of a group of proteins which regulate normal cellular proliferation and
 CC suppression of cellular transformation. Inactivation of the p53 gene has
 CC been implicated in the formation, or progression of a wide variety of
 CC carcinoma. Polypeptides containing at least amino acids 13-41 of p53, or
 CC the DNA encoding these, may be used to inhibit the growth of tumour cells
 CC containing MDM2 gene amplification. (Updated on 25-MAR-2003 to correct PN
 CC field.)
 XX
 XX Sequence 491 AA;
 SQ
 Query Match 73.6%; Score 159; DB 2; Length 491;
 Best Local Similarity 100.0%; Pred. No. 8.3e-147;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 58 QYIMTKELYDEKQKHIVYCSNDLLGDLFGVPSPSVKEHRKIYTIYRNLVVYNQESSDS 117
 DB 59 QYIMTKELYDEKQKHIVYCSNDLLGDLFGVPSPSVKEHRKIYTIYRNLVVYNQESSDS 118
 QY 118 GTSVSENRCHEGSDQKDLVQELQEKPSSSHLSVSRPSTSSRRRAISETEENSDELSCGE 177
 DB 119 GTSVSENRCHEGSDQKDLVQELQEKPSSSHLSVSRPSTSSRRRAISETEENSDELSCGE 178
 QY 178 RQRKHKSIDSISLFDLSLALCVIREICCCRSSSSSESTG 216
 DB 179 RQRKHKSIDSISLFDLSLALCVIREICCCRSSSSSESTG 217
 RESULT 10
 AAR07887
 ID AAR07887 standard; protein; 491 AA.
 XX
 AC AAR07887;
 XX
 XX 25-MAR-2003 (revised)
 DT 28-JAN-1997 (first entry)
 XX
 XX Human MDM-2, involved in tumour-development.
 DE
 XX p53; MDM-2; binding-inhibitor; identification; tumour; cancer; neoplasia;
 KW antibody fusion protein; therapy.
 KW
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 DT 16-OCT-2003 (revised)

DT 01-NOV-1995 (first entry)
 XX Human MDM2 protein.
 DE
 XX MDM2; sarcoma; diagnostic; DNA probe.
 KW
 XX Homo sapiens; (cell line CaCo-2).
 OS
 XX US5420263-A.
 PN
 XX 30-MAY-1995.
 PD
 XX 07-APR-1993; 93US-00044619.
 PF
 XX 07-APR-1992; 92US-00867840.
 PR 23-JUN-1992; 92US-00903103.
 PR
 XX (UYJO) UNIV JOHNS HOPKINS.
 PA
 XX Vogelstein B, Kinzler KW;
 PI WPI; 1995-206312/27.
 XX N-PSDB; AAR94589.
 DR
 XX New human MDM2 cDNA - used to develop prods. for use in the diagnosis and
 PT treatment of tumours.
 PT
 XX Claim 1; Col 23-26; 34pp; English.
 PS
 XX The human MDM2 gene is genetically altered (i.e. amplified) in human
 CC tumour cells. The human MDM2 protein binds to human p53 and allows the
 CC cell to escape from p53-regulated growth. Detecting that the gene has
 CC become amplified or detecting increased gene product expression (using
 CC probes, proteins, antibodies and inhibitors) allows diagnosis and therapy
 CC of cancers such as colorectal carcinoma, lung cancer and chronic
 CC myelogenous leukaemia. (Updated on 15-OCT-2003 to standardise OS field)
 CC
 XX Sequence 491 AA;
 SQ
 Query Match 73.6%; Score 159; DB 2; Length 491;
 Best Local Similarity 100.0%; Pred. No. 8.3e-147;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 58 QYIMTKELYDEKQKHIVYCSNDLLGDLFGVPSPSVKEHRKIYTIYRNLVVYNQESSDS 117
 DB 59 QYIMTKELYDEKQKHIVYCSNDLLGDLFGVPSPSVKEHRKIYTIYRNLVVYNQESSDS 118
 QY 118 GTSVSENRCHEGSDQKDLVQELQEKPSSSHLSVSRPSTSSRRRAISETEENSDELSCGE 177
 DB 119 GTSVSENRCHEGSDQKDLVQELQEKPSSSHLSVSRPSTSSRRRAISETEENSDELSCGE 178
 QY 178 RQRKHKSIDSISLFDLSLALCVIREICCCRSSSSSESTG 216
 DB 179 RQRKHKSIDSISLFDLSLALCVIREICCCRSSSSSESTG 217
 RESULT 10
 AAR07887
 ID AAR07887 standard; protein; 491 AA.
 XX
 AC AAR07887;
 XX
 XX 25-MAR-2003 (revised)
 DT 28-JAN-1997 (first entry)
 XX
 XX Human MDM-2, involved in tumour-development.
 DE
 XX p53; MDM-2; binding-inhibitor; identification; tumour; cancer; neoplasia;
 KW antibody fusion protein; therapy.
 KW
 XX Homo sapiens.
 OS
 XX
 XX Key Location/Qualifiers
 DT 16-OCT-2003 (revised)

```
FT Modified-site 166..169
FT /label= phosphorylation_site
FT /note= "potential casein kinase II phosphorylation site"
FT Binding-site 181..185
FT /label= nuclear_localisation_signal
FT Modified-site 192..195
FT /label= phosphorylation_site
FT /note= "potential casein kinase II phosphorylation site"
FT Domain 223..274
FT /label= acid_activation_domain
FT Modified-site 269..272
FT /label= phosphorylation_site
FT /note= "potential casein kinase II phosphorylation site"
FT Modified-site 290..293
FT /label= phosphorylation_site
FT /note= "potential casein kinase II phosphorylation site"
FT Domain 305..322
FT /label= metal_binding_site
FT Domain 461..478
FT /label= metal_binding_site
XX
XX US5550023-A.
XX
XX 27-AUG-1996.
XX
XX 18-MAY-1994; 94US-00245500.
XX
XX 07-APR-1992; 92US-00867840.
XX 23-JUN-1992; 92US-00903103.
XX 07-APR-1993; 93US-00044619.
XX
XX (UYJO ) UNIV JOHNS HOPKINS.
XX
XX Vogelstein B, Kinzler KW;
XX WPI; 1996-401591/40.
XX N-PSDB; AAT45151.
XX
XX Identification of cpds. interfering with human MDM2/p53 binding - useful
XX as therapeutic agents to treat human neoplastic cells.
XX
XX Claim 26; Col 25-28; 36pp; English.
XX
XX AA07887 represents human MDM-2 derived from a human colon carcinoma cell
XX line, CaCo-2, cDNA library. The MDM-2 protein is used in a method for
XX identifying compounds that interfere with the binding of p53 and MDM-2.
XX In binding the p53 protein, the MDM-2 protein releases a cell from p53-
XX regulated growth, allowing cancers to develop. Therefore compounds
XX identified as interfering with the binding of MDM-2 to p53 are
XX potentially useful in the treatment of human neoplastic cells. In the
XX method pref. one or both of the proteins is a fusion protein esp. with an
XX antibody or antibody fragment which aids separation and identification.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 491 AA;
XX
XX Query Match 73.6%; Score 159; DB 2; Length 491;
XX Best Local Similarity 100.0%; Pred. No. 8.3e-147;
XX Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPSVKHEHKIYTMIRNLVYVNOQESSDS 117
XX 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPSVKHEHKIYTMIRNLVYVNOQESSDS 118
XX
XX 118 GTSVSENCHLEGGSDQDLVQELQEPSSSHLVSRPSTSSRRRAISETEENSDELSGE 177
XX 119 GTSVSENCHLEGGSDQDLVQELQEPSSSHLVSRPSTSSRRRAISETEENSDELSGE 178
XX
XX 178 RQRKHKSDSISLSPDESIALCVIREICCCSSSESTG 216
XX 179 RQRKHKSDSISLSPDESIALCVIREICCCSSSESTG 217
XX
```

```
RESULT 11
AAW15463
ID AAW15463 standard; protein; 491 AA.
XX
XX AAW15463;
XX
XX 25-MAR-2003 (revised)
XX 18-JUN-1997 (first entry)
XX
XX Human MDM2.
XX
XX Human; MDM2 protein; antibody; detection; cancer; diagnosis;
XX p53-regulated growth.
XX
XX Homo sapiens.
XX
XX US5618921-A.
XX
XX 08-APR-1997.
XX
XX 17-FEB-1995; 95US-00390479.
XX
XX 07-APR-1992; 92US-00867840.
XX 23-JUN-1992; 92US-00903103.
XX 07-APR-1993; 93US-00044619.
XX
XX (UYJO ) UNIV JOHNS HOPKINS.
XX
XX Vogelstein B, Kinzler KW, Burrell M, Hill DE;
XX WPI; 1997-225474/20.
XX N-PSDB; AAT66410.
XX
XX Antibodies specific for human MDM2 protein - for diagnosis of cancer.
XX
XX Claim 1; Col 19-24; 35pp; English.
XX
XX This sequence represents the human MDM2 protein. Antibodies that
XX specifically bind to human MDM2 protein may be used for detecting
XX elevated expression of the MDM2 gene in a human tissue or body fluid
XX sample, esp. for cancer diagnosis. The antibodies may be used to
XX interfere with the binding of p53 to MDM2. Elevated levels of MDM2 appear
XX to sequester p53 and allow the cell to escape from p53-regulated growth.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 491 AA;
XX
XX Query Match 73.6%; Score 159; DB 2; Length 491;
XX Best Local Similarity 100.0%; Pred. No. 8.3e-147;
XX Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPSVKHEHKIYTMIRNLVYVNOQESSDS 117
XX 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPSVKHEHKIYTMIRNLVYVNOQESSDS 118
XX
XX 118 GTSVSENCHLEGGSDQDLVQELQEPSSSHLVSRPSTSSRRRAISETEENSDELSGE 177
XX 119 GTSVSENCHLEGGSDQDLVQELQEPSSSHLVSRPSTSSRRRAISETEENSDELSGE 178
XX
XX 178 RQRKHKSDSISLSPDESIALCVIREICCCSSSESTG 216
XX 179 RQRKHKSDSISLSPDESIALCVIREICCCSSSESTG 217
XX
XX RESULT 12
AAW13380
ID AAW13380 standard; protein; 491 AA.
XX
XX AAW13380;
XX
XX 25-MAR-2003 (revised)
XX 05-JUN-1997 (first entry)
XX
```


DE Human MDM2 protein.
 XX Human; MDM2; CaCo-2; colonic; carcinoma; probe; detection; amplification;
 XX elevation; expression; diagnosis; neoplasia; neoplastic transformation;
 KW sarcoma; colorectal; lung cancer; chronic myelogenous leukaemia.
 XX
 OS Homo sapiens.
 XX
 XX US5606044-A.
 XX
 XX 25-FEB-1997.
 XX
 XX 17-FEB-1995; 95US-00390546.
 XX
 XX 07-APR-1992; 92US-00867840.
 XX 23-JUN-1992; 92US-00903103.
 XX 07-APR-1993; 93US-00044619.
 XX
 XX (UWJO) UNIV JOHNS HOPKINS.
 XX
 XX Kinzler KW, Vogelstein B, Hill DE, Burrell M;
 XX
 XX WPI; 1997-153623/14.
 XX N-PSDB; AAT62065.
 XX
 XX Detection of amplification of human MDM2 gene - useful for diagnosis of
 XX neoplasia or potential neoplastic transformation.
 XX
 XX Example 1; Col 21-24; 35pp; English.
 XX
 XX The present sequence is the human MDM2 protein, the cDNA for which was
 XX isolated from a human CaCo-2 colonic carcinoma cell cDNA library using a
 XX murine MDM2 cDNA probe. The MDM2 cDNA can be used as a probe to detect
 XX the amplification or elevated expression of a human MDM2 gene, which is
 XX diagnostic of neoplasia or the potential for neoplastic transformation,
 XX useful for the detection of, e.g. sarcomas, colorectal carcinoma, lung
 XX cancer and chronic myelogenous leukaemia. (Updated on 25-MAR-2003 to
 XX correct PF field.)
 XX
 XX Sequence 491 AA;
 XX
 XX Query Match 73.6%; Score 159; DB 2; Length 491;
 XX Best Local Similarity 100.0%; Pred. No. 8.3e-147;
 XX Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKHKKIYTMIRNLVNVNQESSDS 117
 DB 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKHKKIYTMIRNLVNVNQESSDS 118
 QY 118 GTSVSENRCHEGSGDQKOLVQELQEKPSSSHLSVSPSTSSRRRAISETEENSDELSCGE 177
 DB 119 GTSVSENRCHEGSGDQKOLVQELQEKPSSSHLSVSPSTSSRRRAISETEENSDELSCGE 178
 QY 178 RQRXHKSDSISLSPDESALCVIREICCRSSSSSESTG 216
 DB 179 RQRXHKSDSISLSPDESALCVIREICCRSSSSSESTG 217
 RESULT 13
 AAW13600
 ID AAW13600 standard; protein; 491 AA.
 XX
 XX AAW13600;
 XX
 XX 16-JAN-1998 (first entry)
 XX
 XX Murine double minute 2 protein sequence.
 XX
 XX Mouse; Mdm2; murine double minute; phosphoprotein; binding; modulation;
 KW tumour suppressor; p53; oncogene; cell cycle arrest; p107; antagonist;
 KW inhibition; transcription factor; adenocarcinoma; colon; cancer; breast;
 KW lung; stomach; myeloid leukaemia; lymphoma; hyperproliferative;
 KW restenosis.

XX Mus musculus.
 OS
 XX WO9709343-A2.
 XX
 XX 13-MAR-1997.
 PD
 XX 02-SEP-1996; 96WO-FR001340.
 XX
 XX 04-SEP-1995; 95FR-00010331.
 XX
 XX (RHON) RHONE POULENC RORER SA.
 XX (INRM) INST NAT SANTE & RECH MEDICALE.
 PA
 XX
 XX Tocque B, Dubs-Poterszman M, Wasylyk B;
 PI
 XX WPI; 1997-192837/17.
 XX N-PSDB; AAT61637.
 DR
 XX
 XX Treating cancer with antagonist of oncogenic activity of protein Mdm2 -
 PT or nucleic acid encoding an antagonist, also viral vectors contg. this
 PT nucleic acid.
 XX
 XX Claim 2; Page 26-30; 43pp; French.
 PS
 XX
 XX This is the amino acid sequence of the mouse Mdm2 (murine double minute-
 CC 2) protein, a 90 kD phosphoprotein which binds and modulates the activity
 CC of the tumour suppressor protein p53. It has now been shown that the mdm2
 CC protein itself has oncogenic properties, especially in a p53-null
 CC background. Mdm2 is observed to unblock cell cycle arrest in G1 caused by
 CC over-expression of the p107 protein. This is especially done by the
 CC region covering amino acid 1-134. The invention therefore relates to
 CC antagonists able to inhibit the oncogenic activity of mdm2. These include
 CC fragments of the p53 protein, especially amino acids 1-52, 1-41, 6-41, 16
 CC -25 or 18-23 (AAW13602-6), or fragments of transcription factors e.g.
 CC TF11, TBP or TAF250, which bind amino acids 1-134 of mdm2. Other
 CC inhibitors include compounds which disrupt binding to region 135-491 of
 CC mdm2, e.g. Rb, L5 or the transcription factor E2F. The antagonists are
 CC used to treat e.g. adenocarcinoma of the colon; cancer of the breast,
 CC lung or stomach; myeloid leukaemia; B cell lymphoma, or other
 CC hyperproliferative conditions such as restenosis
 XX
 XX Sequence 491 AA;
 XX
 XX Query Match 73.6%; Score 159; DB 2; Length 491;
 XX Best Local Similarity 100.0%; Pred. No. 8.3e-147;
 XX Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKHKKIYTMIRNLVNVNQESSDS 117
 DB 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKHKKIYTMIRNLVNVNQESSDS 118
 QY 118 GTSVSENRCHEGSGDQKOLVQELQEKPSSSHLSVSPSTSSRRRAISETEENSDELSCGE 177
 DB 119 GTSVSENRCHEGSGDQKOLVQELQEKPSSSHLSVSPSTSSRRRAISETEENSDELSCGE 178
 QY 178 RQRXHKSDSISLSPDESALCVIREICCRSSSSSESTG 216
 DB 179 RQRXHKSDSISLSPDESALCVIREICCRSSSSSESTG 217
 RESULT 14
 AAW48241
 ID AAW48241 standard; protein; 491 AA.
 XX
 XX AAW48241;
 XX
 XX 18-JUN-1998 (first entry)
 XX
 XX Human MDM2.
 XX
 XX Human; MDM2; hMDM2; tumour; cancer; diagnosis; neoplastic disease;
 KW sarcoma; liposarcoma; malignant fibrous histiocytoma; osteosarcoma.

XX OS Homo sapiens.
 XX PN US5736338-A.
 XX PD 07-APR-1998.
 XX PF 17-FEB-1995; 95US-00390517.
 XX PR 07-APR-1992; 92US-00867840.
 XX PR 23-JUN-1992; 92US-00903103.
 XX PR 07-APR-1993; 93US-00044619.
 XX PA (UYJO) UNIV JOHNS HOPKINS.
 XX PI Vogelstein B, Kinzler KW, Burrell M;
 XX DR WPI: 1998-239206/21.
 XX DR N-PSDB; AAV20349.
 XX PT Cancer diagnosis - by determination of MDM2 protein.
 XX PS Claim 1; Col 25-28; 35pp; English.
 XX CC The present sequence represents human MDM2 (hMDM2) which is used in the
 CC method of the present invention. The present invention describes a method
 CC for diagnosing a neoplastic disease caused by overexpression of MDM2
 CC protein. The method comprises detecting an elevated cellular amount of
 CC this protein. The method is useful for the diagnosis of sarcoma,
 CC especially liposarcoma, malignant fibrous histiocytoma or osteosarcoma
 XX SQ Sequence 491 AA;

Query Match 73.6%; Score 159; DB 2; Length 491;
 Best Local Similarity 100.0%; Pred. No. 8.3e-147;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFVSKHKKIYTMIRNLVWVWQESSDS 117
 DB 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFVSKHKKIYTMIRNLVWVWQESSDS 118
 QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSETENSDELSGE 177
 DB 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSETENSDELSGE 178
 QY 178 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 216
 DB 179 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 217

RESULT 15
 ID AAW57241 standard; protein; 491 AA.
 XX AC AAW57241;
 XX DT 10-AUG-1998 (first entry)
 XX DE Human MDM2 protein.
 XX KW Human; p53; MDM2; tumour; growth inhibition; amplification;
 XX KW malignant fibrous histiocytoma; liposarcoma.
 XX OS Homo sapiens.
 XX PN US5756455-A.
 XX PD 26-MAY-1998.
 XX PF 17-FEB-1995; 95US-00390515.
 XX PR 07-APR-1992; 92US-00867840.
 XX PR 23-JUN-1992; 92US-00903103.

PR 07-APR-1993; 93US-00044619.
 XX PA (UYJO) UNIV JOHNS HOPKINS.
 XX PI Vogelstein B, Kinzler KW;
 XX DR WPI: 1998-321574/28.
 XX DR N-PSDB; AAV28876.
 XX PT Inhibiting growth of tumour cells having MDM2 gene amplification - with
 XX PT p53 protein fragment.
 XX PS Claim 1; Col 23-28; 40pp; English.
 XX CC A method has been developed for inhibiting the growth of tumour cells
 CC containing a human MDM2 gene amplification. The method comprises treating
 CC the tumour cells with a DNA molecule that expresses a polypeptide capable
 CC of binding to human MDM2 protein. The present sequence represents human
 CC MDM2 protein. The present invention describes three preferred
 CC polypeptides for binding human MDM2: (1) the polypeptide comprises amino
 CC acids 1-50 of p53 (see AAW57240); (2) the polypeptide comprises amino
 CC acids 13-41 of p53 (see AAW57240); and at least none additional p53
 CC residues on the N- or C-terminal side, provided that the polypeptide
 CC lacks the homooligomerisation domain of p53; (3) the polypeptide
 CC comprises amino acids 13-41 of p53 (see AAW57241) and at least nine
 CC additional p53 residues on the N- or C-terminal side, provided that the
 CC polypeptide lacks amino acids 138-393 of p53. Some malignant fibrous
 CC histiocytomas and liposarcomas have an MDM2 gene amplification, so
 CC detection of increased expression of MDM2 gene products indicates
 CC tumourigenesis
 XX SQ Sequence 491 AA;

Query Match 73.6%; Score 159; DB 2; Length 491;
 Best Local Similarity 100.0%; Pred. No. 8.3e-147;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFVSKHKKIYTMIRNLVWVWQESSDS 117
 DB 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFVSKHKKIYTMIRNLVWVWQESSDS 118
 QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSETENSDELSGE 177
 DB 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISSETENSDELSGE 178
 QY 178 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 216
 DB 179 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 217

Search completed: March 14, 2004, 21:52:53
 Job time : 53.8828 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:51:29 ; Search time 23.0897 Seconds
(without alignments)

482.953 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 216

Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIREICERSSESSTG 216

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	216	100.0	216	4	US-09-510-252-4
2	159	73.6	491	1	US-07-903-103-2
3	159	73.6	491	1	US-08-044-619A-2
4	159	73.6	491	1	US-08-283-911-2
5	159	73.6	491	1	US-08-245-500A-3
6	159	73.6	491	1	US-08-390-546-3
7	159	73.6	491	1	US-08-390-479A-3
8	159	73.6	491	1	US-08-557-393-3
9	159	73.6	491	1	US-08-390-516C-3
10	159	73.6	491	1	US-08-390-517A-3
11	159	73.6	491	1	US-08-390-515A-3
12	159	73.6	491	2	US-08-801-718-3
13	159	73.6	491	4	US-08-170-159A-3
14	159	73.6	491	4	US-08-480-718-44
15	130	60.2	188	4	US-07-603-052-4
16	29	13.4	489	1	US-07-903-103-4
17	29	13.4	489	1	US-08-044-619A-4
18	29	13.4	489	1	US-08-283-911-4
19	29	13.4	489	1	US-08-245-500A-5
20	29	13.4	489	1	US-08-390-546-5
21	29	13.4	489	1	US-08-390-479A-5
22	29	13.4	489	1	US-08-557-393-5
23	29	13.4	489	1	US-08-390-516C-5
24	29	13.4	489	1	US-08-390-517A-5
25	29	13.4	489	1	US-08-390-515A-5
26	29	13.4	489	2	US-08-801-718-5
27	29	13.4	489	4	US-09-170-159A-5

```

28 13.4 489 4 US-09-480-718-46 Sequence 46, Appl
29 6.5 15 1 US-08-277-660A-25 Sequence 25, Appl
30 6.5 15 1 US-08-424-957-10 Sequence 10, Appl
31 6.5 15 3 US-09-035-686-10 Sequence 10, Appl
32 3.7 547 4 US-09-252-991A-27640 Sequence 27640, A
33 3.2 8 4 US-09-428-082B-523 Sequence 523, App
34 3.2 12 4 US-09-428-082B-517 Sequence 517, App
35 3.2 12 4 US-09-428-082B-520 Sequence 520, App
36 3.2 12 4 US-09-428-082B-587 Sequence 587, App
37 3.2 12 4 US-09-428-082B-589 Sequence 589, App
38 3.2 13 4 US-09-428-082B-586 Sequence 586, App
39 3.2 18 4 US-10-083-889-15 Sequence 15, Appl
40 3.2 42 4 US-09-483-847-223 Sequence 223, App
41 3.2 56 4 US-09-489-847-376 Sequence 376, App
42 3.2 66 4 US-09-107-532A-4703 Sequence 4703, Ap
43 3.2 133 4 US-09-252-991A-23380 Sequence 23380, A
44 3.2 136 4 US-09-252-991A-22734 Sequence 22734, A
45 3.2 137 4 US-09-489-039A-13167 Sequence 13167, A

```

ALIGNMENTS

RESULT 1

```

US-09-510-252-4
; Sequence 4, Application US/09510252
; Patent No. 6372490
; GENERAL INFORMATION:
; APPLICANT: Nardabalan, Krishnan
; APPLICANT: Yang, Weijia
; APPLICANT: Schulz, Vincent
; APPLICANT: Curagen Corporation
; TITLE OF INVENTION: MDM INTERACTING PROTEIN AND METHODS OF USE THEREOF
; FILE REFERENCE: 15966-524 MDM US
; CURRENT APPLICATION NUMBER: US/09/510,252
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: USSN 60/121,192
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: USSN 60/122,643
; PRIOR FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-510-252-4

```

```

Query Match 100.0%; Score 216; DB 4; Length 216;
Best Local Similarity 100.0%; Pred. No. 7.6e-198;
Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKELLLKLLKSVGAOKDTYTMKEVLFYLOVI 60
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKELLLKLLKSVGAOKDTYTMKEVLFYLOVI 60
QY 61 MKRLYDEKQOHVVCNSNDLLGLDGLFQVPSFVSKEHKIYTMIRNLVNVNQESSDSGTS 120
DB 61 MKRLYDEKQOHVVCNSNDLLGLDGLFQVPSFVSKEHKIYTMIRNLVNVNQESSDSGTS 120
QY 121 VSENRCHLEGGSDQKDLVQELQEKFPSSSHLVSRPSTSSRRRAISETENSDELSEGEROR 180
DB 121 VSENRCHLEGGSDQKDLVQELQEKFPSSSHLVSRPSTSSRRRAISETENSDELSEGEROR 180
QY 181 KRHKSDSISLSPDESIALCVIREICERSSESSTG 216
DB 181 KRHKSDSISLSPDESIALCVIREICERSSESSTG 216

```

RESULT 2

```

US-07-903-103-2
; Sequence 2, Application US/07903103
; Patent No. 541860

```

```

;
; GENERAL INFORMATION:
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G ST., N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/903,103
; FILING DATE: 07-APR-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/867,840
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.40148
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-903-103-2

Query Match 73.6%; Score 159; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVNVNQESSDS 117
DB 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVNVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISETEENSDLSGE 177
DB 119 GTSVSENCHLEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISETEENSDLSGE 178
QY 178 RQRKHKSDSISLSPDESIALCVIRICCCRSSSESTG 216
DB 179 RQRKHKSDSISLSPDESIALCVIRICCCRSSSESTG 217

RESULT 3
US-08-044-619A-2
; Sequence 2, Application US/08044619A
; Patent No. 542063
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G ST., N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/044,619A
; FILING DATE: 07-APR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: US 07/903,103
; FILING DATE: 23-JUN-1992
; APPLICATION NUMBER: US 07/867,840
; FILING DATE: 07-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.40148
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-044-619A-2

Query Match 73.6%; Score 159; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVNVNQESSDS 117
DB 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVNVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISETEENSDLSGE 177
DB 119 GTSVSENCHLEGGSDQKDLVQLQELQEKPSHLSVRPSTSSRRRAISETEENSDLSGE 178
QY 178 RQRKHKSDSISLSPDESIALCVIRICCCRSSSESTG 216
DB 179 RQRKHKSDSISLSPDESIALCVIRICCCRSSSESTG 217

RESULT 4
US-08-283-911-2
; Sequence 2, Application US/08283911
; Patent No. 551918
; GENERAL INFORMATION:
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G ST., N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001-4597
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/283,911

```

```
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/903,103
/ FILING DATE: 23-JUN-1992
/ APPLICATION NUMBER: US 07/867,840
/ FILING DATE: 07-APR-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KAGAN, SARAH A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 01107.40148
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-508-9100
/ TELEFAX: 202-508-9299
/ TELEX: 197430 BMB UT
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 491 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-245-500A-3

Query Match 73.6%; Score 159; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPSVKEHRKIYTMIRNLVNVNQESSDS 117
DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPSVKEHRKIYTMIRNLVNVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 177
DB 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKRKSDSISLSPDESIALCVIREICCRSSSSSESTG 216
DB 179 RQRKRKSDSISLSPDESIALCVIREICCRSSSSSESTG 217

RESULT 6
US-08-390-546-3
/ Sequence 3, Application US/08390546
/ Patent No. 5606044
/ GENERAL INFORMATION:
/ APPLICANT: BURRELL, MARILEE
/ APPLICANT: HILL, DAVID E.
/ APPLICANT: KINZLER, KENNETH W.
/ APPLICANT: VOGELSTEIN, BERT
/ TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
/ TITLE OF INVENTION: HUMAN TUMORS
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
/ STREET: 1001 G STREET, N.W.
/ CITY: WASHINGTON
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20001
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/390,546
/ FILING DATE: 07-APR-1993
/ CLASSIFICATION: 536
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KAGAN, SARAH A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 01107.42798
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-508-9100
/ TELEFAX: 202-508-9299
/ TELEX: 197430 BMB UT
/ INFORMATION FOR SEQ ID NO: 3:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 491 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-390-546-3

Query Match 73.6%; Score 159; DB 1; Length 491;

/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/903,103
/ FILING DATE: 23-JUN-1992
/ APPLICATION NUMBER: US 07/867,840
/ FILING DATE: 07-APR-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KAGAN, SARAH A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 01107.40148
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-508-9100
/ TELEFAX: 202-508-9299
/ TELEX: 197430 BMB UT
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 491 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-245-500A-3

Query Match 73.6%; Score 159; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPSVKEHRKIYTMIRNLVNVNQESSDS 117
DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPSVKEHRKIYTMIRNLVNVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 177
DB 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKRKSDSISLSPDESIALCVIREICCRSSSSSESTG 216
DB 179 RQRKRKSDSISLSPDESIALCVIREICCRSSSSSESTG 217

RESULT 5
US-08-245-500A-3
/ Sequence 3, Application US/08245500A
/ Patent No. 5550023
/ GENERAL INFORMATION:
/ APPLICANT: BURRELL, MARILEE
/ APPLICANT: HILL, DAVID E.
/ APPLICANT: KINZLER, KENNETH W.
/ APPLICANT: VOGELSTEIN, BERT
/ TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
/ TITLE OF INVENTION: HUMAN TUMORS
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
/ STREET: 1001 G STREET, N.W.
/ CITY: WASHINGTON
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20001
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/245,500A
/ FILING DATE: 07-APR-1993
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KAGAN, SARAH A.
/ REGISTRATION NUMBER: 32,141
/ REFERENCE/DOCKET NUMBER: 01107.42798
/ TELECOMMUNICATION INFORMATION:
```

Best Local Similarity 100.0%; Pred. No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFVSKEHRKIYTMIRNLVWVWQESSDS 117
DB 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFVSKEHRKIYTMIRNLVWVWQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQBEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
DB 119 GTSVSENCHLEGGSDQKDLVQELQBEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 216
DB 179 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 217

RESULT 7
US-08-390-479A-3
; Sequence 3, Application US/08390479A
; Patent No. 5618921
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER & WITCOFF, LTD.
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,479A
; FILING DATE: 02-FEB-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.48992
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-390-479A-3

Query Match 73.6%; Score 159; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFVSKEHRKIYTMIRNLVWVWQESSDS 117
DB 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFVSKEHRKIYTMIRNLVWVWQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQBEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
DB 119 GTSVSENCHLEGGSDQKDLVQELQBEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 216

DB 179 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 217

RESULT 8
US-08-557-393-3
; Sequence 3, Application US/08557393
; Patent No. 5702903
; GENERAL INFORMATION:
; APPLICANT: BURRELL, MARILEE
; APPLICANT: HILL, DAVID E.
; APPLICANT: KINZLER, KENNETH W.
; APPLICANT: VOGELSTEIN, BERT
; TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
; STREET: 1001 G STREET, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,393
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,500
; FILING DATE: 18-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: KAGAN, SARAH A.
; REGISTRATION NUMBER: 32,141
; REFERENCE/DOCKET NUMBER: 01107.42798
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-508-9100
; TELEFAX: 202-508-9299
; TELEX: 197430 BBMB UT
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 491 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-557-393-3

Query Match 73.6%; Score 159; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFVSKEHRKIYTMIRNLVWVWQESSDS 117
DB 59 QYIMTKRLYDEKQOHIVYCSNDLLGDLFGVPSFVSKEHRKIYTMIRNLVWVWQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQBEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 177
DB 119 GTSVSENCHLEGGSDQKDLVQELQBEKPSSSHLSVRPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 216
DB 179 RQRKHKSDSISLSPDESIALCVIREICCRSSSSSESTG 217

RESULT 9
US-08-390-516C-3
; Sequence 3, Application US/08390516C
; Patent No. 5708136
; GENERAL INFORMATION:

APPLICANT: BURRELL, MARILEE
APPLICANT: HILL, DAVID E.
APPLICANT: KINZLER, KENNETH W.
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
TITLE OF INVENTION: HUMAN TUMORS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,516C
FILING DATE: 07-APR-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-516C-3

Query Match 73.6%; Score 159; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 QYIMTKRLYDEKQHHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVNVNQESSDS 117
Db 59 QYIMTKRLYDEKQHHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVNVNQESSDS 118
Qy 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 177
Db 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 178
Qy 178 RQRKHKSDSISLFDSEALCVIREICCRSSSESTG 216
Db 179 RQRKHKSDSISLFDSEALCVIREICCRSSSESTG 217

RESULT 10
US-08-390-517A-3
Sequence 3, Application US/08390517A
Patent No. 5736338
GENERAL INFORMATION:
APPLICANT: BURRELL, MARILEE
APPLICANT: HILL, DAVID E.
APPLICANT: KINZLER, KENNETH W.
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
TITLE OF INVENTION: HUMAN TUMORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G STREET, N.W.
CITY: WASHINGTON
STATE: D.C.

COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,517A
FILING DATE: 07-APR-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-517A-3

Query Match 73.6%; Score 159; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 QYIMTKRLYDEKQHHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVNVNQESSDS 117
Db 59 QYIMTKRLYDEKQHHIVYCSNDLLGDLFGVPFSVKEHRKIYTMIRNLVNVNQESSDS 118
Qy 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 177
Db 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 178
Qy 178 RQRKHKSDSISLFDSEALCVIREICCRSSSESTG 216
Db 179 RQRKHKSDSISLFDSEALCVIREICCRSSSESTG 217

RESULT 11
US-08-390-515A-3
Sequence 3, Application US/08390515A
Patent No. 5756455
GENERAL INFORMATION:
APPLICANT: BURRELL, MARILEE
APPLICANT: HILL, DAVID E.
APPLICANT: KINZLER, KENNETH W.
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
TITLE OF INVENTION: HUMAN TUMORS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,515A
FILING DATE: 07-APR-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-515A-3

Query Match
Best Local Similarity 73.6%; Score 159; DB 1; Length 491;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 117
DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 118

QY 118 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 177
DB 119 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 178

QY 178 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 216
DB 179 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 217

RESULT 12
US-08-801-718-3
Sequence 3, Application US/08801718
Patent No. 5858976
GENERAL INFORMATION:
APPLICANT: BURRELL, MARILEE
APPLICANT: HILL, DAVID E.
APPLICANT: KINZLER, KENNETH W.
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,718
FILING DATE: 14-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/390,515
FILING DATE: 07-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-515A-3

Query Match
Best Local Similarity 73.6%; Score 159; DB 1; Length 491;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 117
DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 118

QY 118 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 177
DB 119 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 178

QY 178 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 216
DB 179 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 217

NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-801-718-3

Query Match
Best Local Similarity 73.6%; Score 159; DB 2; Length 491;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 117
DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 118

QY 118 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 177
DB 119 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 178

QY 178 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 216
DB 179 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 217

RESULT 13
US-09-170-159A-3
Sequence 3, Application US/09170159A
Patent No. 6399755
GENERAL INFORMATION:
APPLICANT: BURRELL, MARILEE
APPLICANT: HILL, DAVID E.
APPLICANT: KINZLER, KENNETH W.
APPLICANT: VOGELSTEIN, BERT
TITLE OF INVENTION: AMPLIFICATION OF HUMAN MDM2 GENE IN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BANNER, BIRCH, MCKIE AND BECKETT
STREET: 1001 G STREET, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/170,159A
FILING DATE: 13-Oct-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: KAGAN, SARAH A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42798
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-508-9100
TELEFAX: 202-508-9299
TELEX: 197430 BMB UT
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 491 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-170-159A-3

Query Match
Best Local Similarity 73.6%; Score 159; DB 4; Length 491;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 117
DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 118

QY 118 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 177
DB 119 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 178

QY 178 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 216
DB 179 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 217

US-09-170-159A-3

Query Match
Best Local Similarity 73.6%; Score 159; DB 4; Length 491;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 117
DB 59 QYIMTKRLYDEKQKHIVYCSNDLLGLDFGVPSFVKHEHKIYTMIVRNLVVYNQESSDS 118

QY 118 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 177
DB 119 GTSVSENRCHEGSDQKDLVQLQEKPSSSHLYVRPSTSSRRRAISETEENSDELSCGE 178

QY 178 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 216
DB 179 RQRKRKSDSISLSPFDESALCVIREICCRSSSSSESTG 217

Db 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPSVKHKRIYTMIRNLVVVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 177
Db 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKHKSIDSLSLFDSESLALCVIREICCRSSSSSESTG 216
Db 179 RQRKHKSIDSLSLFDSESLALCVIREICCRSSSSSESTG 217

RESULT 14

US-09-480-718-44
; Sequence 44, Application US/09480718
; Patent No. 6407062
; GENERAL INFORMATION:
; APPLICANT: Sherr, Charles J
; APPLICANT: Quelle, Dawn E
; APPLICANT: Weber, Jason D.
; APPLICANT: Roussel, Martine F.
; APPLICANT: Frederique, Zindy
; TITLE OF INVENTION: ARF-19, A NOVEL REGULATOR OF THE MAMMALIAN CELL CYCLE
; FILE REFERENCE: 1340-1-023 CIP 1
; CURRENT APPLICATION NUMBER: US/09/480,718
; CURRENT FILING DATE: 2000-01-07
; EARLIER APPLICATION NUMBER: 09/129,855
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 44
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-480-718-44

Query Match 73.6%; Score 159; DB 4; Length 491;
Best Local Similarity 100.0%; Pred.No. 2.8e-143;
Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPSVKHKRIYTMIRNLVVVNQESSDS 117
Db 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPSVKHKRIYTMIRNLVVVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 177
Db 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKHKSIDSLSLFDSESLALCVIREICCRSSSSSESTG 216
Db 179 RQRKHKSIDSLSLFDSESLALCVIREICCRSSSSSESTG 217

RESULT 15

US-09-603-052-4
; Sequence 4, Application US/09603052
; Patent No. 6432116
; GENERAL INFORMATION:
; APPLICANT: Chene, Patrick
; APPLICANT: Hochkeppel, Heinz-Kurt
; TITLE OF INVENTION: Assay for identifying inhibitors of the interaction
; TITLE OF INVENTION: between proteins p53 and dm2
; FILE REFERENCE: MEMB26.001C1
; CURRENT APPLICATION NUMBER: US/09/603,052
; CURRENT FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: EP 95810576.9
; PRIOR FILING DATE: 1995-09-18
; PRIOR APPLICATION NUMBER: PCT/EP96/03957
; PRIOR FILING DATE: 1996-09-10
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 4
; LENGTH: 188

; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-603-052-4
Query Match 60.2%; Score 130; DB 4; Length 188;
Best Local Similarity 100.0%; Pred.No. 5e-116;
Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPSVKHKRIYTMIRNLVVVNQESSDS 117
Db 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSPSVKHKRIYTMIRNLVVVNQESSDS 118
QY 118 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 177
Db 119 GTSVSENCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGE 178
QY 178 RQRKHKSIDS 187
Db 179 RQRKHKSIDS 188

Search completed: March 14, 2004, 21:56:26
Job time : 23.0897 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 19:09:12 ; Search time 89.3793 Seconds
(without alignments)
682.823 Million cell updates/sec

Title: US-10-057-510-4
Perfect score: 1095
Sequence: 1 MONTMSVPTDGAVTTSQIP.....ALCVIREICCRSSSSSESTG 216

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1095	100.0	216	AAB08846	AAB08846 A human M
2	1084.5	99.0	284	AAR75494	AAR75494 Human dou
3	1084.5	99.0	284	AAR75397	AAR75397 Human dou
4	1084.5	99.0	491	AAR42175	AAR42175 Human MDM
5	1084.5	99.0	491	AAR76696	AAR76696 Human MDM
6	1084.5	99.0	491	AAR07887	AAR07887 Human MDM
7	1084.5	99.0	491	AAR15463	AAR15463 Human MDM
8	1084.5	99.0	491	AAR13380	AAR13380 Human MDM
9	1084.5	99.0	491	AAR13600	AAR13600 Murine do
10	1084.5	99.0	491	AAR48241	AAR48241 Human MDM
11	1084.5	99.0	491	AAR57241	AAR57241 Human MDM
12	1084.5	99.0	491	AAR42879	AAR42879 Amino aci
13	1084.5	99.0	491	AAR42971	AAR42971 Amino aci
14	1084.5	99.0	491	AAR94304	AAR94304 Human MDM
15	1084.5	99.0	491	AAR96567	AAR96567 MDM2 onco
16	1084.5	99.0	491	AAR48284	AAR48284 Human MDM
17	1084.5	99.0	491	AAR22654	AAR22654 Human Rin
18	1084.5	99.0	491	AAR22698	AAR22698 Human Rin
19	1084.5	99.0	491	AAR25913	AAR25913 Human dou
20	1084.5	99.0	491	AAR015376	AAR015376 Human Dm2
21	1084.5	99.0	491	AAR21815	AAR21815 Human mdm
22	1084.5	99.0	491	AAR61562	AAR61562 Human Pro
23	1084.5	99.0	1171	AAR32421	AAR32421 Novel hum
24	808	73.8	227	AAR75495	AAR75495 Human dou
25	808	73.8	227	AAR75398	AAR75398 Human dou

26	808	73.8	434	2	AAR75496	Human dou
27	808	73.8	434	2	AAR75399	Human dou
28	805	73.5	489	7	ADD21816	Mouse mdm
29	805	73.5	489	7	ADE61560	Rat Prote
30	804.5	73.5	489	2	AAR42176	Murine MD
31	804.5	73.5	489	2	AAR76697	Mouse MDM
32	804.5	73.5	489	2	AAR07888	Murine MD
33	804.5	73.5	489	2	AAW15464	Murine MD
34	804.5	73.5	489	2	AAW48242	Mouse MDM
35	804.5	73.5	489	2	AAW57246	Mouse MDM
36	804.5	73.5	489	2	AAW42997	Amino aci
37	804.5	73.5	489	2	AAW42972	Amino aci
38	804.5	73.5	489	2	AAW94305	Mouse MDM
39	804.5	73.5	489	5	AAE25914	Mouse dou
40	804.5	73.5	489	5	ABB57099	Mouse lsc
41	804.5	73.5	489	5	AAO15375	Mouse Dm2
42	378	34.5	205	4	ABG29958	Novel hum
43	297.5	27.2	59	7	ADC22294	Protein b
44	278.5	25.4	361	4	ABG29956	Novel hum
45	275	25.1	489	2	AAW10206	Mouse p53

ALIGNMENTS

RESULT 1
AAB08846
ID AAB08846 standard; protein; 216 AA.
XX AC AAB08846;
XX AC
DT 02-JAN-2001 (first entry)
XX
DE A human MDMIP-binding MDM2 polypeptide fragment.
XX

Human; MDM2 interacting polypeptide; MDMIP; MDM2; cell cycle progression;
cell differentiation; cancer; sarcoma; glioma; squamous cell carcinoma;
breast cancer; astrocytoma; leukemia; lymphoma; tumorigenesis;
gene therapy.

XX Homo sapiens.

XX Key Location/Qualifiers
FT Misc-difference 58
FT Misc-difference /note= "Gln encoded by GGCCAG"

XX WO200050590-A1.

XX 31-AUG-2000.

XX 23-FEB-2000; 2000WO-US004582.

XX 23-FEB-1999; 99US-0121192P.

XX 03-MAR-1999; 99US-0122643P.

XX 22-FEB-2000; 2000US-00510252.

XX (CURA-) CURAGEN CORP.

XX Nandabalan K, Yang M, Schulz VP;

XX WPI; 2000-558398/51.

XX N-PSDB; AAA75042.

XX Novel MDM2 interacting protein useful for treating or preventing

XX disorders involving aberrant levels of MDM2 and/or MDM-interacting

XX proteins, comprises a specific amino acid sequence.

XX Claim 11; Fig 2B; 78pp; English.

XX The present sequence represents a fragment of a human MDM2 polypeptide,

XX which binds to a human MDM2 interacting polypeptide (MDMIP). The protein

XX fragment was used as bait in a yeast two hybrid system to identify MDMIP.

XX The MDMIP polypeptide is useful for detecting and removing MDM2

CC polypeptides in a biological sample by forming MDM2-MDMIP complexes.
 CC MDMIP and MDM2 are useful to identify compounds or other agents which
 CC modulate the activity of MDM2 and/or MDMIP-mediated processes. Agents
 CC that modulate the function of MDMIP/MDM2 complexes are useful for
 CC treating and preventing a disease or disorder involving aberrant levels
 CC of MDM2 or MDMIP. MDMIP is also useful for treating diseases caused by
 CC aberrant levels of expression of MDM2 genes, such as disorders of cell
 CC cycle progression, cell differentiation, and transcriptional control,
 CC including cancers such as human sarcoma, glioma, squamous cell carcinoma,
 CC breast cancer, astrocytoma, leukemia and lymphoma, and tumorigenesis.
 CC MDMIP and MDM2 nucleic acids are useful in gene therapy
 XX

Query Match 100.0%; Score 1095; DB 3; Length 216;
 Best Local Similarity 100.0%; Pred. No. 1.4e-106;
 Matches 216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MCNTNMSVPTDGA VTTTQIPASEQETLVPRPKPLLLKLLKSVGAQKDTYTMKEVLFYLYI 60
 DB 1 MCNTNMSVPTDGA VTTTQIPASEQETLVPRPKPLLLKLLKSVGAQKDTYTMKEVLFYLYI 60
 QY 61 MTKELYDEKQOHIVYCSNDLLGLDFGVPSFVKHKKIYTMIRNLVNVNQESSDSGT 120
 DB 61 MTKELYDEKQOHIVYCSNDLLGLDFGVPSFVKHKKIYTMIRNLVNVNQESSDSGT 120
 QY 121 VSENRCHEGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISETENSDELGERQ 180
 DB 121 VSENRCHEGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISETENSDELGERQ 180
 QY 181 KRKSDSISLFDLSALCVIREICCRSSSESTG 216
 DB 181 KRKSDSISLFDLSALCVIREICCRSSSESTG 216

RESULT 2
 AAR75494
 ID AAR75494 standard; protein; 284 AA.
 AC AAR75494;
 XX
 DT 02-FEB-1996 (first entry)
 XX
 DE Human double minute 2 (hdm-2) antibody-binding region fragment 1.
 XX
 KW Human double minute gene 2; hdm-2; antibody binding region; antigen;
 KW cancer; sarcoma; rhabdomyosarcoma; diagnosis; immunoassay.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..284
 FT /note= "amino acids 1-284 of hdm-2 gene product"
 XX
 FN DE4339533-A1.
 XX
 PD 14-JUN-1995.
 XX
 PP 19-NOV-1993; 93DE-04339533.
 XX
 PR 19-NOV-1993; 93DE-04339533.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 PI Zentgraf H, Klein R, Frey M, Martens R;
 XX
 DR WPI; 1995-216248/29.
 DR N-PSDB; AAQ92515.
 XX
 PT Detection of human double minute gene 2 (hdm-2) antibodies - by
 PT incubation with new hdm-2 or antibody-binding hdm-2 fragments; useful in
 PT the detection of specific cancers.
 XX

PS Claim 11; Fig 1; 12pp; German.
 XX
 CC Fragments contg. amino acids 1-284, 58-284 and 58-491 of the hdm-2 (human
 CC double minute 2) gene product are claimed. The overlapping protein
 CC fragments contain binding regions for hdm-2- specific antibodies and are
 CC useful for identifying such antibodies in a claimed immunoassay method.
 CC The presence of anti-hdm-2 antibodies is diagnostic of certain forms of
 CC cancer, e.g. rhabdomyosarcoma
 XX
 SQ Sequence 284 AA;

Query Match 99.0%; Score 1084.5; DB 2; Length 284;
 Best Local Similarity 99.5%; Pred. No. 2.6e-105;
 Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MCNTNMSVPTDGA VTTTQIPASEQETLVPRPKPLLLKLLKSVGAQKDTYTMKEVLFYLYI 59
 DB 1 MCNTNMSVPTDGA VTTTQIPASEQETLVPRPKPLLLKLLKSVGAQKDTYTMKEVLFYLYI 60
 QY 60 IMTKELYDEKQOHIVYCSNDLLGLDFGVPSFVKHKKIYTMIRNLVNVNQESSDSGT 119
 DB 61 IMTKELYDEKQOHIVYCSNDLLGLDFGVPSFVKHKKIYTMIRNLVNVNQESSDSGT 120
 QY 120 VSENRCHEGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISETENSDELGERQ 179
 DB 121 VSENRCHEGSDQKDLVQELQEKPSHLSVRPSTSSRRRAISETENSDELGERQ 180
 QY 180 KRKSDSISLFDLSALCVIREICCRSSSESTG 216
 DB 181 KRKSDSISLFDLSALCVIREICCRSSSESTG 217

RESULT 3
 AAR75397
 ID AAR75397 standard; protein; 284 AA.
 XX
 AC AAR75397;
 XX
 DT 25-MAR-2003 (revised)
 DT 25-JAN-1996 (first entry)
 XX
 DE Human double minute 2 (hdm-2) antibody-binding region fragment 1.
 XX
 KW Human double minute gene 2; hdm-2; antibody binding region; antigen;
 KW cancer; sarcoma; rhabdomyosarcoma; diagnosis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Region 1..284
 FT /note= "amino acids 1-284 of hdm-2 gene product"
 XX
 FN DE4345249-A1.
 XX
 PD 24-MAY-1995.
 XX
 PP 19-NOV-1993; 93DE-04345249.
 XX
 PR 19-NOV-1993; 93DE-04339533.
 XX
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
 XX
 PI Zentgraf H, Klein R, Frey M, Martens R;
 XX
 DR WPI; 1995-195167/26.
 DR N-PSDB; AAQ97261.
 XX
 PT New hdm-2 fragments contg. antibody binding region - used to detect
 PT specific antibodies for diagnosis of cancers, also new DNA sequences
 PT encoding them.
 XX
 PS Claim 2; Fig 1; 11pp; German.
 XX

CC Fragments contg. amino acids 1-284, 58-284 and 58-491 of the hdm-2 (human
CC double minute 2) gene product are claimed. The overlapping protein
CC fragments contain binding regions for hdm-2- specific antibodies and are
CC useful for identifying such antibodies. The presence of anti-hdm-2
CC antibodies is diagnostic of certain forms of cancer, e.g.
CC rhabdomyosarcoma. (Updated on 25-MAR-2003 to correct PF field.)
XX
SQ Sequence 284 AA;

Query Match 99.0%; Score 1084.5; DB 2; Length 284;
Best Local Similarity 99.5%; Pred. No. 2.6e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MCNTNMSVPTDGA VTTT SQIPASEQETLV RPKP LLLK LKSVGAQKDTY TMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGA VTTT SQIPASEQETLV RPKP LLLK LKSVGAQKDTY TMKEVLFYLQY 60
QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 120
QY 120 SVSENRCHEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENSDLSGERQ 179
DB 121 SVSENRCHEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENSDLSGERQ 180
QY 180 RKRHKS DSI SLSPDES LALCVIREIC CERS SSESSTG 216
DB 181 RKRHKS DSI SLSPDES LALCVIREIC CERS SSESSTG 217

RESULT 4
AAR42175
ID AAR42175 standard; protein; 491 AA.
XX
AC AAR42175;
XX
DT 25-MAR-2003 (revised)
DT 05-MAY-1994 (first entry)
XX
DE Human MDM2.
XX
KW p53 gene; tumour suppressor gene; regulation; cellular proliferation;
KW cellular transformation; carcinoma; human; tumour; MDM2; inhibition;
KW gene amplification.

OS Homo sapiens.
XX
PN WO9320238-A2.
XX
PD 14-OCT-1993.
XX
PF 07-APR-1993; 93WO-US003199.
XX
PR 07-APR-1992; 92US-00867840.
PR 23-JUN-1992; 92US-00903103.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Burrell M, Hill DE, Kinzler KW, Vogelstein B;
XX
XX WPI; 1993-336944/42.
DR N-PSDB; AAQ49891.
XX
PT Diagnosing neoplasia from amplification of MDM2 gene - or elevated gene
PT expression, also new DNA, MDM2 protein, antibodies and treatment of
PT sarcoma by inhibiting MDM2 expression.
XX
PS Claim 19; Fig 1; 75pp; English.

CC This sequence is encoded by the MDM2 gene. Amplification of the MDM2 gene
CC is diagnostic of neoplasia or the potential for neoplasia. The protein
CC encoded by this gene interacts with the product of the p53 gene. p53 is a
CC tumour suppressor gene and encodes a protein which appears to be a member

CC of a group of proteins which regulate normal cellular proliferation and
CC suppression of cellular transformation. Inactivation of the p53 gene has
CC been implicated in the formation, or progression of a wide variety of
CC carcinoma. Polypeptides containing at least amino acids 13-41 of p53, or
CC the DNA encoding these, may be used to inhibit the growth of tumour cells
CC containing MDM2 gene amplification. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
SQ Sequence 491 AA;

Query Match 99.0%; Score 1084.5; DB 2; Length 491;
Best Local Similarity 99.5%; Pred. No. 5.8e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 MCNTNMSVPTDGA VTTT SQIPASEQETLV RPKP LLLK LKSVGAQKDTY TMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGA VTTT SQIPASEQETLV RPKP LLLK LKSVGAQKDTY TMKEVLFYLQY 60
QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPSPFSVKEHRKIYTMIRNLYVNVNQESSDSGT 120
QY 120 SVSENRCHEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENSDLSGERQ 179
DB 121 SVSENRCHEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETENSDLSGERQ 180
QY 180 RKRHKS DSI SLSPDES LALCVIREIC CERS SSESSTG 216
DB 181 RKRHKS DSI SLSPDES LALCVIREIC CERS SSESSTG 217

RESULT 5
AAR76696
ID AAR76696 standard; protein; 491 AA.
XX
AC AAR76696;
XX
DT 16-OCT-2003 (revised)
DT 01-NOV-1995 (first entry)
XX
DE Human MDM2 protein.
XX
KW MDM2; sarcoma; diagnostic; DNA probe.
XX
OS Homo sapiens; (cell line CaCo-2).
XX
PN US5420263-A.
XX
PD 30-MAY-1995.
XX
PF 07-APR-1993; 93US-00044619.
XX
PR 07-APR-1992; 92US-00867840.
PR 23-JUN-1992; 92US-00903103.
XX
PA (UYJO) UNIV JOHNS HOPKINS.
XX
PI Vogelstein B, Kinzler KW;
XX
XX WPI; 1995-206312/27.
DR N-PSDB; AAQ94589.
XX
PT New human MDM2 cDNA - used to develop prods. for use in the diagnosis and
PT treatment of tumours.
XX
PS Claim 1; Col 23-26; 34pp; English.

CC The human MDM2 gene is genetically altered (i.e. amplified) in human
CC tumour cells. The human MDM2 protein binds to human p53 and allows the
CC cell to escape from p53-regulated growth. Detecting that the gene has
CC become amplified or detected increased gene product expression (using
CC probes, proteins, antibodies and inhibitors) allows diagnosis and therapy
CC of cancers such as colorectal carcinoma, lung cancer and chronic

```
CC myelogenous leukaemia. (Updated on 16-OCT-2003 to standardise OS field)
XX
SQ Sequence 491 AA;

Query Match          99.0%; Score 1084.5; DB 2; Length 491;
Best Local Similarity 99.5%; Pred. No. 5.8e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYLQY 60

QY 60 IMTKRLYDEKQKHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQKHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDSGT 120

QY 120 SVSENRCHLEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISETEENSDELSGERQ 179
DB 121 SVSENRCHLEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISETEENSDELSGERQ 180

QY 180 RKRHKSDSISLSPDESALCVIREICCRSSSSSESTG 216
DB 181 RKRHKSDSISLSPDESALCVIREICCRSSSSSESTG 217

RESULT 6
AAW07887 standard; protein; 491 AA.
XX
AC AAW07887;
XX
XX 25-MAR-2003 (revised)
DT 28-JAN-1997 (first entry)
XX
DE Human MDK-2, involved in tumour-development.
XX
XX p53; MDM-2; binding-inhibitor; identification; tumour; cancer; neoplasia;
KW antibody fusion protein; therapy.
XX
OS Homo sapiens.
XX
FH Key
FT 166..169 Localisation/Qualifiers
FT Modified-site /label= phosphorylation_site
FT /note= "potential casein kinase II phosphorylation site"
FT Binding-site 181..185
FT /label= nuclear_localisation_signal
FT Modified-site 192..195
FT /label= phosphorylation_site
FT /note= "potential casein kinase II phosphorylation site"
FT Domain 223..274
FT /label= acid_activation_domain
FT Modified-site 269..272
FT /label= phosphorylation_site
FT /note= "potential casein kinase II phosphorylation site"
FT Modified-site 290..293
FT /label= phosphorylation_site
FT /note= "potential casein kinase II phosphorylation site"
FT Domain 305..322
FT /label= metal_binding_site
FT Domain 461..478
FT /label= metal_binding_site
XX
XX US5550023-A.
XX
XX 27-AUG-1996.
XX
XX 18-MAY-1994; 94US-00245500.
XX
XX 07-APR-1992; 92US-00867840.
XX 23-JUN-1992; 92US-00903103.
XX 07-APR-1993; 93US-00044619.
XX
```

```
PA (UYCO ) UNIV JOHNS HOPKINS.
XX
PI Vogelstein B, Kinzler KW;
XX
XX WPI; 1996-401591/40.
DR N-PSDB; AAT45151.
XX
XX Identification of cpds. interfering with human MDM2/p53 binding - useful
XX as therapeutic agents to treat human neoplastic cells.
XX
XX Claim 26; Col 25-28; 36pp; English.
XX
XX AAW07887 represents human MDM-2 derived from a human colon carcinoma cell
XX line, CaCo-2, cDNA library. The MDM-2 protein is used in a method for
XX identifying compounds that interfere with the binding of p53 and MDM-2.
XX In binding the p53 protein, the MDM-2 protein releases a cell from p53-
XX regulated growth, allowing cancers to develop. Therefore compounds
XX identified as interfering with the binding of MDM-2 to p53 are
XX potentially useful in the treatment of human neoplastic cells. In the
XX method pref. one or both of the proteins is a fusion protein esp. with an
XX antibody or antibody fragment which aids separation and identification.
XX (Updated on 25-MAR-2003 to correct PF field.)
XX
XX Sequence 491 AA;
SQ

Query Match          99.0%; Score 1084.5; DB 2; Length 491;
Best Local Similarity 99.5%; Pred. No. 5.8e-105;
Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKDTYTMKEVLFYLQY 60

QY 60 IMTKRLYDEKQKHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDSGT 119
DB 61 IMTKRLYDEKQKHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDSGT 120

QY 120 SVSENRCHLEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISETEENSDELSGERQ 179
DB 121 SVSENRCHLEGGSDQKDLVQELQEKPSSSHLSVRPSTSSRRRAISETEENSDELSGERQ 180

QY 180 RKRHKSDSISLSPDESALCVIREICCRSSSSSESTG 216
DB 181 RKRHKSDSISLSPDESALCVIREICCRSSSSSESTG 217

RESULT 7
AAW15463
ID AAW15463 standard; protein; 491 AA.
XX
XX AAW15463;
XX
XX 25-MAR-2003 (revised)
DT 18-JUN-1997 (first entry)
XX
XX Human MDM2.
XX
XX Human; MDM2 protein; antibody; detection; cancer; diagnosis;
KW p53-regulated growth.
XX
XX Homo sapiens.
XX
XX US5618921-A.
XX
XX 08-APR-1997.
XX
XX 17-FEB-1995; 95US-00390479.
XX
XX 07-APR-1992; 92US-00867840.
XX 23-JUN-1992; 92US-00903103.
XX 07-APR-1993; 93US-00044619.
XX
XX (UYCO ) UNIV JOHNS HOPKINS.
PA
```

XX Vogelstein B, Kinzler KW, Burrell M, Hill DE;
 XX WPI; 1997-225474/20.
 DR N-PSDB; AAT66410.
 XX
 XX Antibodies specific for human MDM2 protein - for diagnosis of cancer.
 XX
 XX Claim 1; Col 19-24; 35pp; English.
 XX
 CC This sequence represents the human MDM2 protein. Antibodies that
 CC specifically bind to human MDM2 protein may be used for detecting
 CC elevated expression of the MDM2 gene in a human tissue or body fluid
 CC sample, esp. for cancer diagnosis. The antibodies may be used to
 CC interfere with the binding of p53 to MDM2. Elevated levels of MDM2 appear
 CC to sequester p53 and allow the cell to escape from p53-regulated growth.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 XX Sequence 491 AA;
 SQ
 Query Match 99.0%; Score 1084.5; DB 2; Length 491;
 Best Local Similarity 99.5%; Pred. No. 5.8e-105;
 Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MCNTNMSVPTDGA VTTTSQIPASEQETLVRPKPLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
 DB 1 MCNTNMSVPTDGA VTTTSQIPASEQETLVRPKPLLKLLKSVGAQKDTYTMKEVLFYLQY 60
 QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSFVKEHRKIYTMIRNLVYVNVQESSDSGT 119
 DB 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSFVKEHRKIYTMIRNLVYVNVQESSDSGT 120
 QY 120 SVSENRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGERQ 179
 DB 121 SVSENRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGERQ 180
 QY 180 RKRHKS DSI SLSPDES LALCVIREICCRSSSSESTG 216
 DB 181 RKRHKS DSI SLSPDES LALCVIREICCRSSSSESTG 217
 RESULT 8
 AAW13380
 ID AAW13380 standard; protein; 491 AA.
 AC AAW13380;
 XX
 XX 25-MAR-2003 (revised)
 DT 05-JUN-1997 (first entry)
 XX
 DE Human MDM2 protein.
 XX
 KW Human; MDM2; CaCo-2; colonic; carcinoma; probe; detection; amplification;
 KW elevation; expression; diagnosis; neoplasia; neoplastic transformation;
 KW sarcoma; colorectal; lung cancer; chronic myelogenous leukaemia.
 XX
 OS Homo sapiens.
 XX
 XX US5606044-A.
 XX
 XX 25-FEB-1997.
 XX
 XX 17-FEB-1995; 95US-00390546.
 XX
 XX 07-APR-1992; 92US-00867840.
 XX
 XX 23-JUN-1992; 92US-00903103.
 XX
 XX 07-APR-1993; 93US-00044619.
 XX
 XX (UYJO) UNIV JOHNS HOPKINS.
 XX
 XX Kinzler KW, Vogelstein B, Hill DE, Burrell M;
 XX WPI; 1997-153623/14.
 DR

DR N-PSDB; AAT62065.
 XX
 XX Detection of amplification of human MDM2 gene - useful for diagnosis of
 PT neoplasia or potential neoplastic transformation.
 XX
 XX Example 1; Col 21-24; 35pp; English.
 XX
 CC The present sequence is the human MDM2 protein, the cDNA for which was
 CC isolated from a human CaCo-2 colonic carcinoma cell cDNA library using a
 CC murine MDM2 cDNA probe. The MDM2 cDNA can be used as a probe to detect
 CC the amplification or elevated expression of a human MDM2 gene, which is
 CC diagnostic of neoplasia or the potential for neoplastic transformation,
 CC useful for the detection of, e.g. sarcomas, colorectal carcinoma, lung
 CC cancer and chronic myelogenous leukaemia. (Updated on 25-MAR-2003 to
 CC correct PF field.)
 XX
 XX Sequence 491 AA;
 SQ
 Query Match 99.0%; Score 1084.5; DB 2; Length 491;
 Best Local Similarity 99.5%; Pred. No. 5.8e-105;
 Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MCNTNMSVPTDGA VTTTSQIPASEQETLVRPKPLLKLLKSVGAQKDTYTMKEVLFYL-QY 59
 DB 1 MCNTNMSVPTDGA VTTTSQIPASEQETLVRPKPLLKLLKSVGAQKDTYTMKEVLFYLQY 60
 QY 60 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSFVKEHRKIYTMIRNLVYVNVQESSDSGT 119
 DB 61 IMTKRLYDEKQOHIVYCSNDLLGDLFGVPFSFVKEHRKIYTMIRNLVYVNVQESSDSGT 120
 QY 120 SVSENRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGERQ 179
 DB 121 SVSENRCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELSGERQ 180
 QY 180 RKRHKS DSI SLSPDES LALCVIREICCRSSSSESTG 216
 DB 181 RKRHKS DSI SLSPDES LALCVIREICCRSSSSESTG 217
 RESULT 9
 AAW13600
 ID AAW13600 standard; protein; 491 AA.
 AC AAW13600;
 XX
 XX 16-JAN-1998 (first entry)
 DT
 XX Murine double minute 2 protein sequence.
 XX
 DE Mouse; Mdm2; murine double minute; phosphoprotein; binding; modulation;
 KW tumour suppressor; p53; oncogene; cell cycle arrest; p107; antagonist;
 KW inhibition; transcription factor; adenocarcinoma; colon; cancer; breast;
 KW lung; stomach; myeloid leukaemia; lymphoma; hyperproliferative;
 KW restenosis.
 XX
 XX Mus musculus.
 XX
 XX W09709343-A2.
 XX
 XX 13-MAR-1997.
 XX
 XX 02-SEP-1996; 96WO-FR001340.
 XX
 XX 04-SEP-1995; 95FR-00010331.
 XX
 XX (RHON) RHONE FOULENC RORER SA.
 XX
 XX (INRM) INST NAT SANTE & RECH MEDICALE.
 XX
 XX Tocque B, Dubs-Poterszman M, Waslyk B;
 XX WPI; 1997-192837/17.
 DR N-PSDB; AAT61637.
 XX

PT Treating cancer with antagonist of oncogenic activity of protein Mdm2 -
 PT or nucleic acid encoding an antagonist, also viral vectors contg. this
 PT nucleic acid.
 XX
 XX
 XX Claim 2; Page 26-30; 43pp; French.
 XX
 CC This is the amino acid sequence of the mouse Mdm2 (murine double minute-
 CC 2) protein, a 90 kD phosphoprotein which binds and modulates the activity
 CC of the tumour suppressor protein p53. It has now been shown that the mdm2
 CC protein itself has oncogenic properties, especially in a p53-null
 CC background. Mdm2 is observed to unblock cell cycle arrest in G1 caused by
 CC over-expression of the p107 protein. This is especially done by the
 CC region covering amino acid 1-134. The invention therefore relates to
 CC antagonists able to inhibit the oncogenic activity of mdm2. These include
 CC fragments of the p53 protein, especially amino acids 1-52, 1-41, 6-41, 16
 CC -25 or 18-23 (AAW13602-6), or fragments of transcription factors e.g.
 CC TP53, TBP or TAP250, which bind amino acids 1-134 of mdm2. Other
 CC inhibitors include compounds which disrupt binding to region 135-491 of
 CC mdm2, e.g. RB, U5 or the transcription factor E2F. The antagonists are
 CC used to treat e.g. adenocarcinoma of the colon; cancer of the breast,
 CC lung or stomach; myeloid leukaemia; B cell lymphoma, or other
 CC hyperproliferative conditions such as restenosis
 XX
 SQ Sequence 491 AA;

Query Match 99.0%; Score 1084.5; DB 2; Length 491;
 Best Local Similarity 99.5%; Pred. No. 5.8e-105;
 Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MCNTNMSVPTDGA VTTTQIPASEQETLVPRPKPLLLKLLKSVGAKQDTYTMKEVLFYL-QY 59
 DB 1 MCNTNMSVPTDGA VTTTQIPASEQETLVPRPKPLLLKLLKSVGAKQDTYTMKEVLFYLQY 60
 QY 60 IMTKRLYDEKQKHIVYCSNDLLGLDGFVPSFVKEHRKIYTMIRNLVVVNQSSSDSGT 119
 DB 61 IMTKRLYDEKQKHIVYCSNDLLGLDGFVPSFVKEHRKIYTMIRNLVVVNQSSSDSGT 120
 QY 120 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGGERQ 179
 DB 121 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGGERQ 180
 QY 180 RKRHKSDSISLFSFDESALCVIREICCRSSSSSESTG 216
 DB 181 RKRHKSDSISLFSFDESALCVIREICCRSSSSSESTG 217

RESULT 10
 AAW48241
 ID AAW48241 standard; protein; 491 AA.
 XX
 AC AAW48241;
 XX
 DT 18-JUN-1998 (first entry)
 XX
 DE Human MDM2.
 XX
 KW Human; MDM2; hMDM2; tumour; cancer; diagnosis; neoplastic disease;
 KW sarcoma; liposarcoma; malignant fibrous histiocytoma; osteosarcoma.
 XX
 OS Homo sapiens.
 XX
 PN US5736338-A.
 XX
 PD 07-APR-1998.
 XX
 PF 17-FEB-1995; 95US-00390517.
 XX
 PR 07-APR-1992; 92US-00867840.
 PR 23-JUN-1992; 92US-00903103.
 PR 07-APR-1993; 93US-00044619.
 XX
 PA (UYJO) UNIV JOHNS HOPKINS.
 XX

PI Vogelstein B, Kinzler KW, Hill DE, Burrell M;
 XX
 DR WPI; 1998-239206/21.
 DR N-PSDB; AAV20549.
 XX
 PT Cancer diagnosis - by determination of MDM2 protein.
 XX
 CC Claim 1; Col 25-28; 35pp; English.
 XX
 CC The present sequence represents human MDM2 (hMDM2) which is used in the
 CC method of the present invention. The present invention describes a method
 CC for diagnosing a neoplastic disease caused by overexpression of MDM2
 CC protein. The method comprises detecting an elevated cellular amount of
 CC this protein. The method is useful for the diagnosis of sarcoma,
 CC especially liposarcoma, malignant fibrous histiocytoma or osteosarcoma
 XX
 SQ Sequence 491 AA;
 Query Match 99.0%; Score 1084.5; DB 2; Length 491;
 Best Local Similarity 99.5%; Pred. No. 5.8e-105;
 Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MCNTNMSVPTDGA VTTTQIPASEQETLVPRPKPLLLKLLKSVGAKQDTYTMKEVLFYL-QY 59
 DB 1 MCNTNMSVPTDGA VTTTQIPASEQETLVPRPKPLLLKLLKSVGAKQDTYTMKEVLFYLQY 60
 QY 60 IMTKRLYDEKQKHIVYCSNDLLGLDGFVPSFVKEHRKIYTMIRNLVVVNQSSSDSGT 119
 DB 61 IMTKRLYDEKQKHIVYCSNDLLGLDGFVPSFVKEHRKIYTMIRNLVVVNQSSSDSGT 120
 QY 120 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGGERQ 179
 DB 121 SVSENRCHLEGGSDOKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGGERQ 180
 QY 180 RKRHKSDSISLFSFDESALCVIREICCRSSSSSESTG 216
 DB 181 RKRHKSDSISLFSFDESALCVIREICCRSSSSSESTG 217

RESULT 11
 AAW57241
 ID AAW57241 standard; protein; 491 AA.
 XX
 AC AAW57241;
 XX
 DT 10-AUG-1998 (first entry)
 XX
 DE Human MDM2 protein.
 XX
 KW Human; p53; MDM2; tumour; growth inhibition; amplification;
 KW malignant fibrous histiocytoma; liposarcoma.
 XX
 OS Homo sapiens.
 XX
 PN US5756455-A.
 XX
 PD 26-MAY-1998.
 XX
 PF 17-FEB-1995; 95US-00390515.
 XX
 PR 07-APR-1992; 92US-00867840.
 PR 23-JUN-1992; 92US-00903103.
 PR 07-APR-1993; 93US-00044619.
 XX
 PA (UYJO) UNIV JOHNS HOPKINS.
 XX
 PI Vogelstein B, Kinzler KW;
 XX
 DR WPI; 1998-321574/28.
 DR N-PSDB; AAV28876.
 XX
 PT Inhibiting growth of tumour cells having MDM2 gene amplification - with
 PT p53 protein fragment.

XX 24-NOV-1999; 99WO-US027907.
 XX PF
 XX 25-NOV-1998; 98US-0109891P.
 XX PR
 XX 17-FEB-1999; 99US-0120549P.
 XX PR
 XX (GENE-) GENETICA INC.
 XX PA
 XX Hannon GJ, Beach DH;
 XX PI
 XX WPI; 2000-400055/34.
 XX DR
 XX N-PSDB; AAA29389.
 XX
 PT New method for increasing the proliferative capacity of cell lines
 PT comprises administering agents reversibly activating telomerase activity
 PT and reversibly inactivating Rb/INK4 and/or p53 pathways useful in
 PT treating age related diseases.
 XX
 XX Claim 5; Page 120; 123pp; English.
 XX PS
 XX
 CC The invention concerns methods and reagents for extending the life-span,
 CC e.g. the number of mitotic divisions, of a cell. The method relies on
 CC activation of a telomerase activity and inhibition of one or both of a
 CC retinoblastoma (Rb)/INK4 pathway or a p53 pathway. Phosphorylation of Rb
 CC by cyclin-dependent kinases, cdk4 and cdk6, releases the cells into the
 CC division cycle. Binding of INK4 family members, e.g. the tumour
 CC suppressor p16INK4a, inhibits kinase activity and results in growth
 CC arrest. Rb inactivators can selectively and reversibly inactivate an
 CC Rb/INK4 pathway, especially an Rb/p16INK4a pathway. The oncoprotein MDM2
 CC is a cellular inhibitor of Rb/E2F function and the p53 tumour suppressor
 CC and can also be used in the methods. Other molecules which can be used
 CC include cdk4 or cdk6 mutants. In particular, a cdk4 mutant is one which
 CC differs from at one or more of residues K22, R24, H95 and/or D97.
 CC Additional constructs include a papilloma virus E7 protein, or other
 CC viral oncoprotein which bypasses Rb and/or p53. Antisense constructs of
 CC the Rb and p16INK4a genes may also be used. The methods are useful for
 CC increasing the proliferative capacity of cells. The cells are
 CC subsequently of use in pharmaceutical and cosmetic preparations used to
 CC treat conditions related to (premature) ageing, e.g. macular degeneration
 CC and arteriosclerosis. The cells can also be used to replace tumour cell
 CC lines in vitro and for studies on biochemical and physiological aspects
 CC of growth and differentiation. Long lived (immortal) cells could also be
 CC of use in the production of normal or genetically engineered
 CC biotechnology products
 XX
 XX Sequence 491 AA;
 XX SQ
 Query Match 99.0%; Score 1084.5; DB 3; Length 491;
 Best Local Similarity 99.5%; Pred. No. 5.8e-105;
 Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDTYTMKEVLFYLQY 59
 DB 1 MCNTNMSVPTDGAVTTSQIPASEQETLVPRPKLLKLLKSVGAQKDTYTMKEVLFYLQY 60
 QY 60 IMTKRLYDEKQOHVYCSNDLLGDLFGVPFSFVKHRKIYTMIRNLVVVNOQESSDSGT 119
 DB 61 IMTKRLYDEKQOHVYCSNDLLGDLFGVPFSFVKHRKIYTMIRNLVVVNOQESSDSGT 120
 QY 120 SVSENKCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGERQ 179
 DB 121 SVSENKCHLEGGSDQKDLVQELQEEKPSSSHLVSRPSTSSRRRAISETEENSDELGERQ 180
 QY 180 RKHKSDSISLSPDESIALCVIREICCRSSSSESTG 216
 DB 181 RKHKSDSISLSPDESIALCVIREICCRSSSSESTG 217

Search completed: March 14, 2004, 21:44:55
 Job time : 91.3793 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:47:53 ; Search time 12.6621 seconds
(without alignments)
888.256 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 216

Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIREICCRSSSSSTG 216

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	159	73.6	491	1 MDM2 HUMAN	Q00987 homo sapien
2	54	25.0	487	1 MDM2 CANFA	P56950 canis famil
3	52	24.1	491	1 MDM2 HORSE	P58951 equus cabal
4	36	16.7	466	1 MDM2 MESAU	Q60524 mesocricetu
5	29	13.4	489	1 MDM2 MOUSE	P23804 mus musculu
6	6	4.6	473	1 MDM2 XENLA	P56273 xenopus lae
7	9	4.2	1159	1 SOR2 MOUSE	Q96pr5 mus musculu
8	8	3.7	445	1 MDM2 BEARE	Q42354 brachydanio
9	8	3.7	641	1 HRPX SESEY	P41501 pseudomonas
10	7	3.2	149	1 FUR HAEDU	P71333 haemophilus
11	7	3.2	150	1 ALI5 HEVER	Q39967 hevea bras
12	7	3.2	173	1 CRAA SPAEH	Q64211 spalax leuc
13	7	3.2	196	1 CRA2 SPAEH	P15990 spalax leuc
14	7	3.2	237	1 PUR7 LISIN	Q92an6 listeria in
15	7	3.2	237	1 PUR7 LISMO	Q8y6b9 listeria mo
16	7	3.2	230	1 SLB2 XENLA	Q9y6p6 xenopus lae
17	7	3.2	271	1 KSGA HELPU	Q9zj17 h dimethyla
18	7	3.2	290	1 TPGS HUMAN	Q9y2b4 homo sapien
19	7	3.2	345	1 GRA4 TOXGO	Q27002 toxoplasma
20	7	3.2	376	1 YBC9 YEAST	P38201 saccharomyc
21	7	3.2	408	1 EIS MYCBO	P59772 mycobacteri
22	7	3.2	408	1 EIS MYCTU	P71727 mycobacteri
23	7	3.2	440	1 MUAI BACHD	Q9x610 bacillus ha
24	7	3.2	458	1 VIMI XENLA	P24789 xenopus lae
25	7	3.2	467	1 DPO4 CORGL	Q8nnp4 corynebacte
26	7	3.2	477	1 GLGA STRPN	Q97qs5 streptococc
27	7	3.2	495	1 ACH3 BOVIN	Q07263 bos taurus
28	7	3.2	496	1 ACH3 CHICK	P09481 gallus gall
29	7	3.2	499	1 ACH3 RAT	P04757 rattus norv
30	7	3.2	503	1 ACH3 HUMAN	P32297 homo sapien
31	7	3.2	511	1 SYT6 MOUSE	Q9r0n8 mus musculu
32	7	3.2	511	1 SYT6 RAT	Q62746 rattus norv
33	7	3.2	526	1 ESR2 PIG	Q9xsw2 sus scrofa

RESULT 1

MDM2_HUMAN

ID_MDM2_HUMAN STANDARD; PRT; 491 AA.
AC Q00987; Q13226; Q13297; Q13298; Q13300; Q13301; Q9UGI3;
AC Q90WT8;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
DE Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (Hdm2).
GN MDM2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM MDM2).
RX MEDLINE=92310576; PubMed=1614537;
RA Oliner J.D., Kinzler K.W., Meltzer P.S., George D.L.,
RA Vogelstein B.;
RT "Amplification of a gene encoding a p53-associated protein in human
RT sarcomas.";
RL Nature 358:80-83(1992).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS MDM2-A; -B; -C; -D AND -E).
RC TISSUE=Ovarian carcinoma;
RX MEDLINE=96313107; PubMed=8705862;
RA Sigalae I., Calvert A.H., Anderson J.J., Neal D.E., Lunec J.;
RT "Alternatively spliced mdm2 transcripts with loss of p53 binding
RT domain sequences: transforming ability and frequent detection in human
RT cancer.";
RL Nat. Med. 2:912-917(1996).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM MDM2-ALPHA).
RX MEDLINE=20065171; PubMed=10597303;
RA Veldhoen N., Metcalfe S., Milner J.;
RT "A novel exon within the mdm2 gene modulates translation initiation in
RT vitro and disrupts the p53-binding domain of mdm2 protein.";
RL Oncogene 18:7026-7033(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM MDM2).
RX MEDLINE=22388257; PubMed=12477932;
RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schachwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM MDM2).
RX MEDLINE=23388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Hopkins R.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

ALIGNMENTS

34 7 3.2 562 1 ILVD PYRAE
35 7 3.2 585 1 CGA2 YEAST
36 7 3.2 585 1 YW04 SCHPO
37 7 3.2 596 1 TF3B YEAST
38 7 3.2 646 1 YFF4 YEAST
39 7 3.2 791 1 KGGG HUMAN
40 7 3.2 802 1 K6A5 HUMAN
41 7 3.2 834 1 SRC1 YEAST
42 7 3.2 863 1 K6A5 MOUSE
43 7 3.2 910 1 DNJM MYCPN
44 7 3.2 1039 1 TPR1 SCHPO
45 7 3.2 1095 1 AT9B HUMAN

Q8ZYU6 pyrobaculum
P38817 saccharomyc
Q9hgm7 schizosacch
P29056 saccharomyc
P43549 saccharomyc
P49619 homo sapien
O75582 homo sapien
Q03707 saccharomyc
Q8C050 mus musculu
P75354 mycoplasma
O42668 schizosacch
O43861 homo sapien

RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Wozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grinnwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krywinski M.I., Skalska U., Smalusz D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [6]
 RP SEQUENCE OF 6-491 FROM N.A. (ISOFORM MDM2-A1).
 RA Liang H., Atkins H., Adel-Fattah R., Suayun R., Lunec J.,
 RT "Genomic organisation of the human MDM2 oncogene and relationship to
 RT its alternatively spliced mRNA's";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE OF 1-24 FROM N.A.
 RX MEDLINE=95380270; PubMed=7651818;
 RA Zauberman A., Flusberg D., Haupt Y., Barak Y., Oren M.,
 RT "A functional p53-responsive intronic promoter is contained within
 RT the human mdm2 gene.";
 RL Nucleic Acids Res. 23:2584-2592(1995).
 RN [8]
 RP SEQUENCE OF 1-9 FROM N.A.
 RX MEDLINE=97413643; PubMed=9270029;
 RA Landers J.E., Cassel S.B., George D.L.,
 RT "Translational enhancement of mdm2 oncogene expression in human tumor
 RT cells containing a stabilized wild-type p53 protein.";
 RL Cancer Res. 57:3562-3568(1997).
 RN [9]
 RP SEQUENCE OF 301-481 FROM N.A.
 RX MEDLINE=20542019; PubMed=11087894;
 RA Taubert H., Kappeler M., Meyer A., Bartel F., Schlott T.,
 RA Lautenschlaeger C., Bache M., Schmidt H., Wuerl P.,
 RT "A MboII polymorphism in exon 11 of the human MDM2 gene occurring in
 RT normal blood donors and in soft tissue sarcoma patients: an
 RT indication for an increased cancer susceptibility?";
 RL Mutat. Res. 456:39-44(2000).
 RN [10]
 RP MUTAGENESIS OF CYS-464.
 RX MEDLINE=98111004; PubMed=9450543;
 RA Honda R., Tanaka H., Yasuda H.,
 RT "Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53.";
 RL FEBS Lett. 420:25-27(1997).
 RN [11]
 RP MUTAGENESIS OF CYS-449.
 RX MEDLINE=20190101; PubMed=10723139;
 RA Honda R., Yasuda H.,
 RT "Activity of MDM2, a ubiquitin ligase, toward p53 or itself is
 RT dependent on the RING finger domain of the ligase.";
 RL Oncogene 19:1473-1476(2000).
 RN [12]
 RP MUTAGENESIS.
 RX MEDLINE=20187618; PubMed=10722742;
 RA Fang S., Jensen J.P., Ludwig R.L., Vousden K.H., Weissman A.M.,
 RT "Mdm2 is a RING finger-dependent ubiquitin protein ligase for itself
 RT and p53.";
 RL J. Biol. Chem. 275:8945-8951(2000).
 RN [13]
 RP MUTAGENESIS OF CYS-441 AND CYS-478.
 RX MEDLINE=20076498; PubMed=10608892;
 RA Sharp D.A., Kratowicz S.A., Sank M.J., George D.L.,
 RT "Stabilization of the MDM2 oncoprotein by interaction with the
 RT structurally related DMX protein.";
 RL J. Biol. Chem. 274:38189-38196(1999).
 RN [14]
 RP NUCLEOLAR LOCALIZATION SIGNAL.
 RX MEDLINE=20173879; PubMed=10707090;
 RA Lohrum M.A.E., Ashcroft M., Kubbutat M.H.G., Vousden K.H.,

RT "Identification of a cryptic nucleolar-localization signal in MDM2.";
 RL Nat. Cell Biol. 2:179-181(2000).
 RN [15]
 RP PHOSPHORYLATION BY ATM.
 RX MEDLINE=20079591; PubMed=10611322;
 RA Khosravi R., Maya R., Gottlieb T., Oren M., Shiloh Y., Shkedy D.,
 RT "Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation
 RT in response to DNA damage";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:14973-14977(1999).
 RN [16]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 25-109 IN COMPLEX WITH P53.
 RX MEDLINE=97081050; PubMed=8875929;
 RA Kussie P.H., Gorina S., Marechal V., Elenbaas B., Moreau J.,
 RA Levine A.J., Pavletich N.P.,
 RT "Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor
 RT transactivation domain.";
 RL Science 274:948-953(1996).
 CC -1- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
 CC APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
 CC FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
 CC TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
 CC TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS.
 CC -1- COFACTOR: zinc is required for ubiquitin ligase E3 activity.
 CC -1- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
 CC SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOBLASTOMA PROTEIN
 CC (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION
 CC FACTOR.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC. EXPRESSED
 CC PREDOMINANTLY IN THE NUCLEOLAR INTERACTION WITH ARF(P14)
 CC RESULTS IN THE LOCALIZATION OF BOTH PROTEINS TO THE NUCLEOLUS. THE
 CC NUCLEOLAR LOCALIZATION SIGNALS IN BOTH ARF(P14) AND MDM2 MAY BE
 CC NECESSARY TO ALLOW EFFICIENT NUCLEOLAR LOCALIZATION OF BOTH
 CC PROTEINS.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=alternative splicing; Named isoforms=8;
 CC Name=Mdm2-Q00987-1; Sequence=Displayed;
 CC IsoId=Q00987-1; Sequence=VSP_003208;
 CC Name=Mdm2-A;
 CC IsoId=Q00987-2; Sequence=VSP_003208;
 CC Name=Mdm2-A1;
 CC IsoId=Q00987-3; Sequence=VSP_003208, VSP_003214;
 CC Name=Mdm2-B;
 CC IsoId=Q00987-4; Sequence=VSP_003209;
 CC Name=Mdm2-C;
 CC IsoId=Q00987-5; Sequence=VSP_003211;
 CC Name=Mdm2-D;
 CC IsoId=Q00987-6; Sequence=VSP_003210;
 CC Name=Mdm2-E;
 CC IsoId=Q00987-7; Sequence=VSP_003212, VSP_003213;
 CC Name=Mdm2-alpha;
 CC IsoId=Q00987-8; Sequence=VSP_003207;
 CC -1- TISSUE SPECIFICITY: UBIQUITOUS. ISOFORMS MDM2-A, -B, -C, -D AND -E
 CC ARE OBSERVED IN A RANGE OF HUMAN CANCERS BUT ABSENT IN NORMAL
 CC TISSUES.
 CC -1- INDUCTION: By DNA damage.
 CC -1- DOMAIN: REGION 1 IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
 CC GL ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
 CC REGION 11 CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
 CC INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
 CC FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
 CC ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
 CC PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
 CC ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
 CC ITSELF.
 CC -1- PTM: PHOSPHORYLATED IN RESPONSE TO IONIZING RADIATION IN AN ATM-
 CC DEPENDENT MANNER.
 CC -1- DISEASE: Seems to be amplified in certain tumors (including soft
 CC tissue sarcomas, osteosarcomas and gliomas). A higher frequency of
 CC splice variants lacking p53 binding domain sequences was found in
 CC late-stage and high-grade ovarian and bladder carcinomas. Four of
 CC the splice variants show loss of p53 binding.
 CC -1- MISCELLANEOUS: MDM2 RING FINGER MUTATIONS THAT FAILED TO
 CC UBIQUITINATE P53 IN VITRO DID NOT TARGET P53 FOR DEGRADATION WHEN

Query Match 73.6%; Score 159; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 7.1e-159;
 Matches 159; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDS 117
 Db 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQESSDS 118

Qy 118 GTSVSENCHLGGSDQKDLVQELQBEKPSHLSVSRPSTSSRRRAISSETENSDLSGE 177
 Db 119 GTSVSENCHLGGSDQKDLVQELQBEKPSHLSVSRPSTSSRRRAISSETENSDLSGE 178

Qy 178 RORKEHKSISLSDESALCVIREICERRSSSESTG 216
 Db 179 RORKEHKSISLSDESALCVIREICERRSSSESTG 217

RESULT 2
 MDM2_CANFA STANDARD; PRT; 487 AA.
 AC P56950; Q95KNS;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
 DE Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (Cdm2).
 GN MDM2.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE OF 1-484 FROM N.A.
 RX MEDLINE=20218666; PubMed=10754200;
 RA Nasir L., Burr P.D., McFarlane S.T., Gault E., Thompson H.,
 RA Argyle D.J.;
 RT "Cloning, sequence analysis and expression of the cDNAs encoding the
 RT canine and equine homologues of the mouse double minute 2 (mdm2)
 RT proto-oncogene.";
 RL Cancer Lett. 152:9-13 (2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS MDM2 AND MDM2-ALPHA).
 RX MEDLINE=20065171; PubMed=10597303;
 RA Veldhoen N., Metcalfe S., Wilner J.;
 RT "A novel exon within the mdm2 gene modulates translation initiation in
 RT vitro and disrupts the p53-binding domain of mdm2 protein.";
 RL Oncogene 18:7026-7033 (1999).
 CC -!- FUNCTION: INHIBITS P53. AND P73-MEDIATED CELL CYCLE ARREST AND
 CC APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
 CC FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
 CC TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
 CC TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS (BY SIMILARITY).
 CC -!- COFACTOR: Zinc is required for ubiquitin ligase E3 activity (By
 CC similarity).
 CC -!- SUBUNIT: BINDS P53, P73, ARF (P14), RIBOSOMAL PROTEIN L5 AND
 CC SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOBLASTOMA PROTEIN
 CC (RB). E1A-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION
 CC FACTOR (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. Expressed
 CC predominantly in the nucleoplasm (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Mdm2;
 CC IsoId=P56950-1; Sequence=Displayed;
 CC Name=Mdm2-alpha;
 CC IsoId=P56950-2; Sequence=VSP_003206;
 CC -!- TISSUE SPECIFICITY: Isoform Mdm2-alpha is present in lymphoid and
 CC testicular tissues.
 CC -!- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
 CC G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
 CC REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
 CC INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC

FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
 ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
 PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
 ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
 ITSELF (BY SIMILARITY).
 -!- SIMILARITY: Belongs to the MDM2 / MDM4 family.
 -!- SIMILARITY: Contains 1 RING-type zinc finger.
 -!- SIMILARITY: Contains 1 RANBP2-type zinc finger.
 -!- SIMILARITY: Contains 1 SWIB domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; AF100705; AAF67833.1; -;
 EMBL; AF322416; AAG42840.1; -;
 DR HSP; Q9UNT8; IYCR.
 DR InterPro; IPR003121; SWIB.
 DR InterPro; IPR001876; Znf_RanGDP.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF02201; SWIB.1.
 DR Pfam; PF00641; zf-RanBP; 1.
 DR SMART; SMO0184; RING; 1.
 DR PROSITE; PS01358; ZF_RANBP2_1; 1.
 DR PROSITE; PS0199; ZF_RANBP2_2; 1.
 DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW Nuclear protein; Ligase; Ub1 conjugation pathway; Zinc; Zinc-finger;
 KW Metal-binding; Alternative splicing.
 FT DOMAIN 27 107 SWIB.
 FT DOMAIN 179 185 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 FT DOMAIN 190 202 NUCLEAR EXPORT SEQUENCE.
 FT DOMAIN 210 304 ARF BINDING.
 FT DOMAIN 210 215 POLY-SER.
 FT DOMAIN 242 331 REGION II.
 FT DOMAIN 243 301 ASP/GLU-RICH (ACIDIC).
 FT ZN_FING 299 328 RANBP2-TYPE.
 FT ZN_FING 434 475 RING-TYPE.
 FT DOMAIN 462 469 NUCLEOLAR LOCALIZATION SIGNAL
 (POTENTIAL).
 FT VARSPPLIC 1 61 Missing (in isoform Mdm2-alpha).
 FT CONFLICT 11 11 /FTId=VSP_003206.
 FT CONFLICT 238 239 G->D (IN REF. 2).
 FT CONFLICT 239 239 QD->HH (IN REF. 2).
 SQ SEQUENCE 487 AA; 54696 MW; 60CDB470A32A8B69 CRC64;
 Query Match 25.0%; Score 54; DB 1; Length 487;
 Best Local Similarity 100.0%; Pred. No. 1.5e-48;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQ 111
 Db 59 QYIMTKRLYDEKQKHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIRNLVNVNQ 112

RESULT 3
 MDM2_HORSE STANDARD; PRT; 491 AA.
 ID MDM2_HORSE
 AC P56951;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
 DE Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (Edm2).
 GN MDM2.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;

RN SEQUENCE FROM N.A.
 RX MEDLINE=2021866; PubMed=10754200;
 RA Nasir L., Burr P.D., McFarlane S.T., Gault E., Thompson H.,
 RA Argyle D.J.;
 RT "Cloning, sequence analysis and expression of the cDNAs encoding the
 RT canine and equine homologues of the mouse double minute 2 (mdm2)
 RL Cancer Lett. 152:9-13(2000).
 CC -1- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
 CC APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN
 CC FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
 CC TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
 CC TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS (BY SIMILARITY).
 CC -1- COPACTOR: Zinc is required for ubiquitin ligase E3 activity (By
 CC similarity).
 CC -1- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
 CC SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOBLASTOMA PROTEIN
 CC (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION
 CC FACTOR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. Expressed
 CC predominantly in the nucleoplasm (By similarity).
 CC -1- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
 CC G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
 CC REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
 CC INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
 CC FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
 CC ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
 CC PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
 CC ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
 CC ITSELF (BY SIMILARITY).
 CC -1- SIMILARITY: Belongs to the MDM2 / MDM4 family.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
 CC -1- SIMILARITY: Contains 1 SWIB domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF121140; AAF28866.1; --
 CC HSPF; Q9UMT8; 1YCR.
 CC InterPro; IPR003121; SWIB.
 CC InterPro; IPR001876; Znf_RanGDP.
 CC InterPro; IPR001841; Znf_ring.
 CC Pfam; PF02201; SWIB; 1.
 CC Pfam; PF00641; zf-RanBP; 1.
 CC SMART; SM00184; RING; 1.
 CC PROSITE; PS01358; ZF_RANBP2_1; 1.
 CC PROSITE; PS00199; ZF_RANBP2_2; 1.
 CC PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
 CC PROSITE; PS50089; ZF_RING_2; 1.
 CC Nuclear protein; Ligase; Ub1 conjugation pathway; Zinc; Zinc-finger;
 CC Metal-binding.
 CC -----
 CC FT DOMAIN 27 107 SWIB.
 CC FT DOMAIN 179 185 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
 CC FT DOMAIN 190 202 NUCLEAR EXPORT SEQUENCE.
 CC FT DOMAIN 210 304 ARF BINDING.
 CC FT DOMAIN 210 315 POLY-SER.
 CC FT DOMAIN 242 331 REGION II.
 CC FT DOMAIN 243 301 ASP/GLU-RICH (ACIDIC).
 CC FT ZN_FING 299 328 RANBP2-TYPE.
 CC FT ZN_FING 438 479 RING-TYPE.
 CC FT DOMAIN 466 473 NUCLEOLAR LOCALIZATION SIGNAL
 CC (POTENTIAL).
 CC SQ SEQUENCE 491 AA; 55279 MW; 641E033D5C1DEC39 CRC64;

Query March 24.1%; Score 52; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 1.9e-46;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 58 QVIMTKRLYDEKQOHIVVCSNDLLGLFGVPSPFSVKEHKIYTMIRNLVVV 109
 DB 59 QVIMTKRLYDEKQOHIVVCSNDLLGLFGVPSPFSVKEHKIYTMIRNLVVV 110
 RESULT 4
 MDM2 MESAU STANDARD; PRT; 466 AA.
 ID MDM2 MESAU
 AC Q60524;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
 Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein) (P-fragment).
 GN MDM2.
 OS Mesocricetus auratus (Golden hamster).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 CC Mesocricetus.
 CC Mesocricetus.
 CC NCBI_TaxID=10036;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=95300112; PubMed=7780969;
 RA Chang K.W., Iaconi S., Mangold K.A., Hubchak S., Scarpelli D.G.;
 RT "Multiple genetic alterations in hamster pancreatic ductal
 RT adenocarcinomas";
 RL Cancer Res. 55:2560-2568(1995).
 CC -1- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
 CC APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
 CC FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
 CC TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
 CC TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS (BY SIMILARITY).
 CC -1- COPACTOR: Zinc is required for ubiquitin ligase E3 activity (By
 CC similarity).
 CC -1- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
 CC SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOBLASTOMA PROTEIN
 CC (RB), E1A-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION
 CC FACTOR (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear and cytoplasmic. Expressed
 CC predominantly in the nucleoplasm (By similarity).
 CC -1- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
 CC G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
 CC REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
 CC INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
 CC FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
 CC ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
 CC PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
 CC ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
 CC ITSELF (BY SIMILARITY).
 CC -1- DISEASE: The gene for this protein is overexpressed in some
 CC tumors.
 CC -1- SIMILARITY: Belongs to the MDM2 / MDM4 family.
 CC -1- SIMILARITY: Contains 1 RING-type zinc finger.
 CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.
 CC -1- SIMILARITY: Contains 1 SWIB domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; U10982; AAC52425.1; --
 CC HSPF; Q9UMT8; 1YCR.
 CC InterPro; IPR003121; SWIB.
 CC InterPro; IPR001876; Znf_RanGDP.
 CC InterPro; IPR001841; Znf_ring.
 CC Pfam; PF02201; SWIB; 1.

DR Pfam; PF00641; zf-RanBP; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS00199; ZF_RANBP2_2; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS00589; ZF_RING_2; 1.
KW Nuclear protein; Ligase; Ub1 conjugation pathway; Proto-oncogene;
KW Zinc; Zinc-finger; Metal-binding.
FT NON_TER 1 1
FT DOMAIN 19 98 SWIB.
FT DOMAIN 169 175 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 176 188 NUCLEAR EXPORT SEQUENCE.
FT DOMAIN 196 201 POLY-SER.
FT DOMAIN 196 290 ARF BINDING.
FT DOMAIN 228 317 REGION II.
FT DOMAIN 209 287 ASP/GLU-RICH (ACIDIC).
FT ZN_FING 285 314 RANBP2-TYPE.
FT ZN_FING 419 460 RING-TYPE.
FT DOMAIN 447 484 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT NON_TER 466 466
SQ SEQUENCE 466 AA; 52390 MW; 78A3042163CF939 CRC64;
Query Match 16.7%; Score 36; DB 1; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 58 QYIMTKRLYDEKQHIYVCSNLLGLDFGVPSFSVK 93
DB 50 QYIMTKRLYDEKQHIYVCSNLLGLDFGVPSFSVK 85
RESULT 5
MDM2_MOUSE
ID MDM2_MOUSE STANDARD; PRT; 489 AA.
AC P23804; Q61040; Q64330;
DT 01-NOV-1991 (Rel. 20, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
DE Mdm2) (Oncoprotein Mdm2) (Double minute 2 protein).
GN MDM2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A. (ISOFORM MDM2-P90).
RX MEDLINE=91224107; PubMed=2026149;
RA Fakhrazadeh S.S., Trusko S.P., George D.L.;
RT "Tumorigenic potential associated with enhanced expression of a gene
RT that is amplified in a mouse tumor cell line."
RL EMBO J. 10:1565-1569 (1991).
[2]
RN SEQUENCE FROM N.A. (ISOFORM MDM2-P90).
RC STRAIN=129/SV;
RX MEDLINE=97074574; PubMed=8917101;
RA Jones S.N., Asari-Lari M.A., Hancock A.R., Jones W.J., Gibbs R.A.,
RA Donohower L.A., Bradley A.;
RT "Genomic organization of the mouse double minute 2 gene."
RL Gene 175:209-213 (1996).
[3]
RN SEQUENCE FROM N.A. (ISOFORM MDM2-P90).
RC STRAIN=129/SV;
RX MEDLINE=96299630; PubMed=8660994;
RA de Oca Luna R.M., Tabor A.D., Eberspaecher H., Hulboy D.L.,
RA Worth L.L., Colman M.S., Finlay C.A., Lozano G.;
RT "The organization and expression of the mdm2 gene."
RL Genomics 33:352-357 (1996).
[4]
RN SEQUENCE FROM N.A. (ISOFORMS MDM2-P90 AND MDM2-P76).
RX MEDLINE=99175199; PubMed=10075719;
RA Saucedo L.J., Myers C.D., Perry M.E.;

RT "Multiple murine double minute gene 2 (MDM2) proteins are induced by
RT ultraviolet light."
RL J. Biol. Chem. 274:8161-8168 (1999).
[5]
RN NUCLEOLAR LOCALIZATION SIGNAL.
RX MEDLINE=20180080; PubMed=10713175;
RA Rousset J.D., Kuo M.-L., Bothner B., DiGiannarino E.L., Kriwacki R.W.,
RA Roussel M.F., Sherr C.J.;
RT "Cooperative signals governing ARF-mdm2 interaction and nucleolar
RT localization of the complex."
RL Mol. Cell. Biol. 20:2517-2528 (2000).
[6]
RN PHOSPHORYLATION BY ATM.
RX MEDLINE=20079591; PubMed=10611322;
RA Khosravi R., Maya R., Gottlieb T., Oren M., Shiloh Y., Shkedy D.;
RT "Rapid ATM-dependent phosphorylation of MDM2 precedes p53 accumulation
RT in response to DNA damage."
RL Proc. Natl. Acad. Sci. U.S.A. 96:14973-14977 (1999).
CC -!- FUNCTION: INHIBITS P53- AND P73-MEDIATED CELL CYCLE ARREST AND
CC APOPTOSIS BY BINDING ITS TRANSCRIPTIONAL ACTIVATION DOMAIN.
CC FUNCTIONS AS A UBIQUITIN LIGASE E3, IN THE PRESENCE OF E1 AND E2,
CC TOWARD P53 AND ITSELF. PERMITS THE NUCLEAR EXPORT OF P53 AND
CC TARGETS IT FOR PROTEASOME-MEDIATED PROTEOLYSIS.
CC -!- COFACTOR: Zinc is required for ubiquitin ligase E3 activity.
CC -!- SUBUNIT: BINDS P53, P73, ARF(P14), RIBOSOMAL PROTEIN L5 AND
CC SPECIFICALLY TO RNA. CAN INTERACT ALSO WITH RETINOBLASTOMA PROTEIN
CC (RB), E2F-ASSOCIATED PROTEIN P300 AND THE E2F1 TRANSCRIPTION
CC FACTOR.
CC -!- SUBCELLULAR LOCATION: NUCLEAR AND CYTOPLASMIC. EXPRESSED
CC PREDOMINANTLY IN THE NUCLEOLUS. INTERACTION WITH ARF(P14)
CC RESULTS IN THE LOCALIZATION OF BOTH PROTEINS TO THE NUCLEOLUS. THE
CC NUCLEOLAR LOCALIZATION SIGNALS IN BOTH ARF(P14) AND MDM2 MAY BE
CC NECESSARY TO ALLOW EFFICIENT NUCLEOLAR LOCALIZATION OF BOTH
CC PROTEINS.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Mdm2-p90;
CC IsoId=P23804-1; Sequence=Displayed;
CC Note=Isoform Mdm2-p76 can also be produced by alternative
CC initiation at Met-50 of isoform Mdm2-p90, but is produced more
CC efficiently by alternative splicing;
CC Name=Mdm2-p76;
CC IsoId=P23804-2; Sequence=VSP_003215;
CC Note=Does not bind to P53;
CC Event=Alternative initiation;
CC Comment=2 isoforms, Mdm2-P90 (shown here) and Mdm2-p76, are
CC produced by alternative initiation at Met-1 and Met-50. Isoform
CC Mdm2-p76 is produced more efficiently by alternative splicing;
CC -!- TISSUE SPECIFICITY: UBIQUITOUSLY EXPRESSED AT LOW-LEVEL THROUGHOUT
CC EMBRYO DEVELOPMENT AND IN ADULT TISSUES. MDM2-P90 IS MUCH MORE
CC ABUNDANT THAN MDM2-P76 IN TESTIS, BRAIN, HEART, AND KIDNEY, BUT IN
CC THE THYMUS, SPLEEN, AND INTESTINE, THE LEVELS OF THE MDM2 PROTEINS
CC ARE ROUGHLY EQUIVALENT.
CC -!- INDUCTION: By UV light.
CC -!- DOMAIN: REGION I IS SUFFICIENT FOR BINDING P53 AND INHIBITING ITS
CC G1 ARREST AND APOPTOSIS FUNCTIONS. IT ALSO BINDS P73 AND E2F1.
CC REGION II CONTAINS MOST OF A CENTRAL ACIDIC REGION REQUIRED FOR
CC INTERACTION WITH RIBOSOMAL PROTEIN L5 AND A PUTATIVE C4-TYPE ZINC
CC FINGER. THE RING FINGER DOMAIN WHICH COORDINATES TWO MOLECULES OF
CC ZINC INTERACTS SPECIFICALLY WITH RNA WHETHER OR NOT ZINC IS
CC PRESENT AND MEDIATES THE HETERO-OLIGOMERIZATION WITH MDM4. IT IS
CC ALSO ESSENTIAL FOR ITS UBIQUITIN LIGASE E3 ACTIVITY TOWARD P53 AND
CC ITSELF.
CC -!- PTM: PHOSPHORYLATED IN RESPONSE TO IONIZING RADIATION IN AN ATM-
CC DEPENDENT MANNER.
CC -!- DISEASE: The gene for this protein is amplified in a mouse tumor
CC cell line.
CC -!- SIMILARITY: Belongs to the MDM2 / MDM4 family.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -!- SIMILARITY: Contains 1 SWIB domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration


```
FT TURN 68 69
FT STRAND 70 72
FT TURN 74 75
FT TURN 77 82
FT HELIX 83 83
FT TURN 86 88
FT STRAND 89 90
FT TURN 92 100
FT HELIX 101 102
FT TURN 103 105
FT STRAND 104 105
SQ SEQUENCE 473 AA; 53464 MW; 7DA668DE8B3BEE01 CRC64;

Query Match 4.6%; Score 10; DB 1; Length 473;
Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 LYDEKQOHIV 74
| | | | |
DB 62 LYDEKQOHIV 71

RESULT 7
SOR2_MOUSE STANDARD; PRT; 1159 AA.
AC Q9E9F5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE VPS10 domain-containing receptor SorCS2 precursor.
GN SORCS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=21095188; PubMed=11165493;
RA Rezaei M., Hermy G., Riedel I.B., Hampe W., Schaller H.C.,
RA Hermans-Borgmeyer I.;
RT Identification of SorCS2, a novel member of the VPS10 domain
RT containing receptor family, prominently expressed in the developing
RT mouse brain.;
RL Mech. Dev. 100:335-338(2001).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Detected in brain, lung, testis and heart.
CC -!- DEVELOPMENTAL STAGE: Expression is highest in developing brain.
CC Transiently expressed in all 3 germ layers.
CC -!- SIMILARITY: Contains 1 PKD domain.
CC -!- SIMILARITY: Contains 6 BNR repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY004316; AAF88135.1; -
CC MGD; MGI:1932289; SorCS2.
CC InterPro; IPR002860; GH_BNR.
CC InterPro; IPR00601; PKD.
CC InterPro; IPR006581; VPS10.
CC Pfam; PF02012; BNR; 6.
CC SMART; SM00089; PKD; 1.
CC SMART; SM00602; VPS10; 1.
CC PROSITE; PS00093; PKD; 1.
CC Signal; Transmembrane; Repeat.
KW SIGNAL 1 49 POTENTIAL.
FT CHAIN 50 1159 VPS10 DOMAIN-CONTAINING RECEPTOR SORCS2.
FT DOMAIN 50 1078 LUMENAL (POTENTIAL).
FT TRANSMEM 1079 1099 POTENTIAL.
```

```
FT DOMAIN 1100 1159 CYTOPLASMIC (POTENTIAL).
FT REPEAT 182 193 BNR 1.
FT REPEAT 232 243 BNR 2.
FT REPEAT 273 284 BNR 3.
FT REPEAT 468 479 BNR 4.
FT REPEAT 545 556 BNR 5.
FT REPEAT 587 598 BNR 6.
FT DOMAIN 786 876 PKD.
FT DOMAIN 715 720 POLY-SER.
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 362 362 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 600 600 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 830 830 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 891 891 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 902 902 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1159 AA; 128872 MW; 9B9F501F2A525D2A CRC64;

Query Match 4.2%; Score 9; DB 1; Length 1159;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 207 ERSSSEST 215
| | | | |
DB 713 ERSSSEST 721

RESULT 8
MDM2_BRARE STANDARD; PRT; 445 AA.
AC O42354;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-protein ligase E3 Mdm2 (EC 6.3.2.-) (p53-binding protein
DE Mdm2) (Double minute 2 protein).
GN MDM2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Neel H., Plette J.;
RA "Partial cDNA nucleotide sequence of the zebrafish homolog of Mdm2.";
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May bind p53 protein and may function as a ubiquitin
CC ligase E3.
CC -!- COFACTOR: Zinc is required for ubiquitin ligase E3 activity (By
CC similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC -!- SIMILARITY: Belongs to the MDM2 / MDM4 family.
CC -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC -!- SIMILARITY: Contains 1 RANBP2-type zinc finger.
CC -!- SIMILARITY: Contains 1 SWIB domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF010255; AAB64176.1; -
CC HSSP; P56273; 1YCQ.
CC ZFIN; ZDB-GENE-990415-153; mdm2.
CC InterPro; IPR003121; SWIB.
CC InterPro; IPR001876; Znf_RanGDP.
CC InterPro; IPR001841; Znf_Ring.
CC Pfam; PF02201; SWIB; 1.
CC SMART; SM00184; RING; 1.
```

```
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS0199; ZF_RANBP2_2; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS0089; ZF_RING_2; 1.
KW Nuclear protein; Ligase; Ubi conjugation pathway; Zinc; Zinc-finger;
KW Metal-binding.
FT DOMAIN 20 100 SWIB.
FT DOMAIN 141 145 POLY-ARG.
FT DOMAIN 160 166 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 171 183 NUCLEAR EXPORT SEQUENCE.
FT DOMAIN 190 279 ARF BINDING.
FT DOMAIN 222 306 REGION II.
FT DOMAIN 220 276 ASP/GLU-RICH (ACIDIC).
FT ZN_FING 274 303 RANBP2-TYPE.
FT ZN_FING 332 433 RING-TYPE.
FT DOMAIN 420 427 NUCLEOLAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 445 AA; 49949 MW; 6FA8175A8A8E6261 CRC64;

Query Match 3.7%; Score 8; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 RKXKSDS 187
DB 162 RKXKSDS 169
|||||

RESULT 9
ID HRPK_PSEY STANDARD; PRT; 641 AA.
AC P41501.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Pathogenicity locus protein hrpK.
GN HRPK.
OS Pseudomonas syringae (pv. syringae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=321;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pss61;
RX MEDLINE=94148760; PubMed=8106313;
RA Xiao Y.; Heu S.; Yi J.; Lu Y.; Hutcheson S.W.;
RT "Identification of a putative alternate sigma factor and
characterization of a multicomponent regulatory cascade controlling
the expression of Pseudomonas syringae pv. syringae Pss61 hrp and
hrmA genes."
RT hrmA Genes."
RL J. Bacteriol. 176:1025-1036 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pss61;
RX MEDLINE=94100578; PubMed=8274770;
RA Heu S.; Hutcheson S.W.;
RT "Nucleotide sequence and properties of the hrmA locus associated with
the Pseudomonas syringae pv. syringae 61 hrp gene cluster."
RT Mol. Plant Microbe Interact. 6:553-564 (1993).
CC
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
-----
CC Query Match 3.7%; Score 8; DB 1; Length 641;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY SEQUENCE 641 AA; 67678 MW; EC098941E5B46C8E CRC64;
DR EMBL; U03855; AAA17653.1; -
SQ SEQUENCE 641 AA; 67678 MW; EC098941E5B46C8E CRC64;
```

```
DR PROSITE; PS01358; ZF_RANBP2_1; 1.
DR PROSITE; PS0199; ZF_RANBP2_2; 1.
DR PROSITE; PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE; PS0089; ZF_RING_2; 1.
KW Nuclear protein; Ligase; Ubi conjugation pathway; Zinc; Zinc-finger;
KW Metal-binding.
FT DOMAIN 20 100 SWIB.
FT DOMAIN 141 145 POLY-ARG.
FT DOMAIN 160 166 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 171 183 NUCLEAR EXPORT SEQUENCE.
FT DOMAIN 190 279 ARF BINDING.
FT DOMAIN 222 306 REGION II.
FT DOMAIN 220 276 ASP/GLU-RICH (ACIDIC).
FT ZN_FING 274 303 RANBP2-TYPE.
FT ZN_FING 332 433 RING-TYPE.
FT DOMAIN 420 427 NUCLEOLAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 445 AA; 49949 MW; 6FA8175A8A8E6261 CRC64;

Query Match 3.7%; Score 8; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 RKXKSDS 187
DB 162 RKXKSDS 169
|||||

RESULT 9
ID HRPK_PSEY STANDARD; PRT; 641 AA.
AC P41501.
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Pathogenicity locus protein hrpK.
GN HRPK.
OS Pseudomonas syringae (pv. syringae).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=321;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pss61;
RX MEDLINE=94148760; PubMed=8106313;
RA Xiao Y.; Heu S.; Yi J.; Lu Y.; Hutcheson S.W.;
RT "Identification of a putative alternate sigma factor and
characterization of a multicomponent regulatory cascade controlling
the expression of Pseudomonas syringae pv. syringae Pss61 hrp and
hrmA genes."
RT hrmA Genes."
RL J. Bacteriol. 176:1025-1036 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Pss61;
RX MEDLINE=94100578; PubMed=8274770;
RA Heu S.; Hutcheson S.W.;
RT "Nucleotide sequence and properties of the hrmA locus associated with
the Pseudomonas syringae pv. syringae 61 hrp gene cluster."
RT Mol. Plant Microbe Interact. 6:553-564 (1993).
CC
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
-----
CC Query Match 3.7%; Score 8; DB 1; Length 641;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY SEQUENCE 641 AA; 67678 MW; EC098941E5B46C8E CRC64;
DR EMBL; U03855; AAA17653.1; -
SQ SEQUENCE 641 AA; 67678 MW; EC098941E5B46C8E CRC64;
```

```
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 RSSSSSEST 215
DB 74 RSSSSSEST 81
|||||

RESULT 10
ID FUR_HAEDU STANDARD; PRT; 149 AA.
AC P71333.
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Ferric uptake regulation protein (Ferric uptake regulator).
GN FUR OR HD0367.
OS Haemophilus ducreyi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=730;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35000HP / ATCC 700724;
RX MEDLINE=97075920; PubMed=8918243;
RA Carson S.D.B.; Thomas C.E.; Elkins C.E.;
RT "Cloning and sequencing of a Haemophilus ducreyi fur homolog."
RL Gene 176:125-129 (1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=35000HP / ATCC 700724;
RA Munson R.S. Jr.; Ray W.C.; Mahairas G.; Sabo P.; Mungur R.;
RA Johnson L.; Nguyen D.; Wang J.; Forst C.; Hood L.;
RT "The complete genome sequence of Haemophilus ducreyi."
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC
-----
CC -!- FUNCTION: Acts as a global negative controlling element, employing
Fe(2+) as a cofactor to bind the operator of the repressed genes
(BY similarity).
CC
-----
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (BY similarity).
CC
-----
CC -!- SIMILARITY: Belongs to the Fur family.
CC
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
-----
CC EMBL; U37224; AAC44583.1; -
CC EMBL; AE017151; AAP95338.1; -
DR PIR; JC5097; JC5097.
DR InterPro; IPR002481; FUR.
DR Pfam; PF01475; FUR; 1.
DR ProDom; PD002003; FUR; 1.
DR Transcription regulation; Repressor; DNA-binding; Iron; Zinc;
KW Complete proteome.
FT DOMAIN 88 93 HIS-RICH.
FT METAL 96 96 ZINC (BY SIMILARITY).
FT METAL 99 99 ZINC (BY SIMILARITY).
SQ SEQUENCE 149 AA; 17404 MW; 754A3198A07C12FA CRC64;

Query Match 3.2%; Score 7; DB 1; Length 149;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 KLLKSVG 42
DB 7 KLLKSVG 13
|||||

RESULT 11
ALUS_HEVER STANDARD; PRT; 150 AA.
ID ALL5_HEVER
```



```

CC of the lens.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=Minor;
CC IsoId=P15990-1; Sequence=Displayed;
CC Name=Major;
CC IsoId=Q64211-1; Sequence=External;
CC -!- SIMILARITY: Belongs to the small heat shock protein (HSP20)
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M17249; AAA66165.1; -.
CC PIR; A28332; A28332.
CC InterPro; IPR001436; Crystallin alpha.
CC InterPro; IPR003090; Crystallin_N.
CC InterPro; IPR002068; HSP20.
CC InterPro; IPR008978; HSP20_chap.
CC Pfam; PF00525; crystallin; 1.
CC Pfam; PF00111; HSP20; 1.
CC PRINTS; PR00299; ACRYSTALLIN.
CC ProDom; PD001193; Crystallin_N; 1.
CC PROSITE; PS01031; HSP20; 1.
CC Eye lens protein; Acetylation; Alternative splicing.
CC MOD_RES 1 ACETYLATION (PROBABLY).
CC FT MOD_RES 1 ACETYLATION (PROBABLY).
CC SQ SEQUENCE 196 AA; 22532 MW; 4D4090D43EB9F906 CRC64;

Query Match 3.2%; Score 7; DB 1; Length 196;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 QEEKPS 148
Db 186 QEEKPS 192
|||||

RESULT 14
PUR7 LISIN STANDARD; PRT; 237 AA.
AC Q92AF6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)
DE (SAICAR synthetase).
GN PURC OR LIN1884.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
"Comparative genomics of Listeria species.";
Science 294:849-852(2001).
CC -!- CATALYTIC ACTIVITY: ATP + 5-amino-1-(5-phospho-D-
CC ribosyl)imidazole-4-carboxylate + L-aspartate = ADP + phosphate +
CC (S)-2-[5-amino-1-(5-phospho-D-ribosyl)imidazole-4-

```

```

CC ribosyl)imidazole-4-carboxylate + L-aspartate = ADP + phosphate +
CC (S)-2-[5-amino-1-(5-phospho-D-ribosyl)imidazole-4-
CC carboxamido]succinate.
CC -!- PATHWAY: De novo purine biosynthesis; seventh step.
CC -!- SIMILARITY: Belongs to the SAICAR synthetase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AL596170; CAC97114.1; -.
CC PIR; AB1668; AB1668.
CC ListList; LINO1884; -.
CC HAMAP; MF_00137; -.
CC InterPro; IPR001636; SAICAR_synt.
CC Pfam; PF01259; SAICAR_synt; 1.
CC ProDom; PD003043; SAICAR_synt; 1.
CC TIGRFAMs; TIGR00081; purC; 1.
CC PROSITE; PS01057; SAICAR_SYNTHETASE 1; 1.
CC PROSITE; PS01058; SAICAR_SYNTHETASE 2; 1.
CC Purine biosynthesis; Ligase; Complete proteome.
CC SQ SEQUENCE 237 AA; 26895 MW; 9D803BAC44A11EF CRC64;

Query Match 3.2%; Score 7; DB 1; Length 237;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 RAISETE 168
Db 71 RAISETE 77
|||||

RESULT 15
PUR7 LISMO STANDARD; PRT; 237 AA.
AC Q8Y6B9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Phosphoribosylaminoimidazole-succinocarboxamide synthase (EC 6.3.2.6)
DE (SAICAR synthetase).
GN PURC OR LMO1772.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EGD-e / Serovar 1/2a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kretz J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schluter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Weiland J., Cossart P.;
"Comparative genomics of Listeria species.";
Science 294:849-852(2001).
CC -!- CATALYTIC ACTIVITY: ATP + 5-amino-1-(5-phospho-D-
CC ribosyl)imidazole-4-carboxylate + L-aspartate = ADP + phosphate +
CC (S)-2-[5-amino-1-(5-phospho-D-ribosyl)imidazole-4-
CC carboxamido]succinate.
CC -!- PATHWAY: De novo purine biosynthesis; seventh step.
CC -!- SIMILARITY: Belongs to the SAICAR synthetase family.
CC

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMEL; AL591981; CAC99850.1; -;
 DR PIR; AD1296; AD1296.
 DR List11st; LMO01772; -;
 DR HAMAP; MF 00137; -; 1.
 DR InterPro; IPR001636; SAICAR synt.
 DR Pfam; PF01259; SAICAR synt; 1.
 DR ProDom; PD001043; SAICAR synt; 1.
 DR TIGRFAMs; TIGR00081; purG; 1.
 DR PROSITE; PS01057; SAICAR SYNTHETASE 1; 1.
 DR PROSITE; PS01058; SAICAR SYNTHETASE_2; 1.
 KW Purine biosynthesis; Ligase; Complete proteome.
 SQ SEQUENCE 237 AA; 26841 MW; BF3G6ED18010425F CRC64;

Query Match 3.2% Score 7; DB 1; Length 237;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 RAISETE 168
 |||||
 Db 71 RAISETE 77

Search completed: March 14, 2004, 21:53:31
 Job time : 14.6621 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 14, 2004, 21:35:23 ; Search time 29.7931 Seconds
(without alignments)
697.389 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 1095

Sequence: 1 MCNTNMSVPTDGAVTTSQIP.....ALCVIREICERSSSSSTG 216

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 78.*

1: Pir1.*

2: Pir2.*

3: Pir3.*

4: Pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1084.5	99.0	491	1 S24354	p53-binding protei
2	804.5	73.5	489	2 S15349	mdm2 protein - mou
3	277	25.3	489	2 S71955	MDM2-like p53-bind
4	103.5	9.5	851	2 A59284	unconventional myo
5	99.5	9.1	2253	2 T30336	nuclear/mitotic ap
6	96.5	8.8	635	1 SYCHAL	5-aminolevulinate
7	93	8.5	365	2 G84559	probable NAM (no a
8	92	8.4	520	2 S35575	myosin heavy chain
9	91.5	8.4	433	2 G64594	hemolysin secretio
10	90.5	8.3	433	2 H71917	methyl-accepting c
11	90	8.2	2037	2 T16881	hypothetical prote
12	89.5	8.2	751	2 T01449	cytoskeletal prote
13	89	8.1	820	2 T04227	hypothetical prote
14	88	8.0	900	2 C71339	probable DNA misa
15	87.5	8.0	572	1 VERULC	lamin C - human
16	87.5	8.0	664	1 VERULA	lamin A - human
17	87.5	8.0	892	2 T40040	GTPase-activator p
18	87.5	8.0	1106	2 T19948	hypothetical prote
19	87.5	8.0	1263	2 T13465	hypothetical prote
20	87	7.9	500	2 S55785	nucleolar protein
21	87	7.9	2278	1 S56274	FAB1 protein - yea
22	86.5	7.9	361	2 T43536	peptidyl prollyl ci
23	86.5	7.9	1417	2 A57570	Bloom's syndrome r
24	86.5	7.9	1535	2 T49042	hypothetical prote
25	86.5	7.9	3788	2 T30851	lysosomal traffick
26	86	7.9	990	2 T51618	nucleolar phosphop
27	86	7.9	1021	2 T51497	hypothetical prote
28	86	7.9	1066	2 G84746	hypothetical prote
29	86	7.9	2326	2 B47447	calcium channel pr

30	85.5	7.8	698	1 S31714	NAD+-protein ADP-r
31	85.5	7.8	1233	2 T14157	serine/threonine p
32	85	7.8	574	2 S04333	lamin C - mouse
33	85	7.8	605	2 T43974	hypothetical prote
34	85	7.8	621	2 T20307	hypothetical prote
35	85	7.8	665	2 S28182	lamin A - mouse
36	85	7.8	672	2 T20310	hypothetical prote
37	85	7.8	862	2 T01798	hypothetical prote
38	85	7.8	967	2 A64710	type III restricti
39	85	7.8	1818	1 S73852	hypothetical prote
40	85	7.8	1876	2 T28627	vitellogenin - Rip
41	85	7.8	1931	2 A59234	slow myosin heavy
42	84.5	7.7	377	1 BWHXC	polysaccharide exp
43	84.5	7.7	462	2 I53414	lamin C2 - mouse
44	84.5	7.7	665	2 S27267	lamin A - rat
45	84.5	7.7	735	2 T45059	hypothetical prote

ALIGNMENTS

RESULT 1

S24354

p53-binding protein mdm2 - human

N:Alternate names: mdm-2 oncogene; mouse double minute 2 homolog; p53-associated phosph

C:Species: Homo sapiens (man)

C>Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000

C:Accession: S24354; S57338; G02026

R:Oliner, J.D.; Kinzler, K.W.; Meltzer, P.S.; George, D.L.; Vogelstein, B.

Nature 358, 80-83, 1992

A:Title: Amplification of a gene encoding a p53-associated protein in human sarcomas.

A:Reference number: S24354; MUID:92310576; PMID:1614537

A:Accession: S24354

A:Molecule type: mRNA

A:Residues: 1-491 <Ofr>

A:Cross-references: EMBL:Z12020; NID:G35211; PIDN:CAA78055.1; PID:G35212

R:Zauberman, A.; Flusberg, D.; Haupt, Y.; Barak, Y.; Oren, M.

Nucleic Acids Res. 23, 2584-2592, 1995

A:Title: A functional p53-responsive intronic promoter is contained within the human md

A:Reference number: S57338; MUID:95380270; PMID:7651818

A:Accession: S57338

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-16, 'P', 18-24 <ZAU>

A:Cross-references: EMBL:U28935; NID:G904033; PIDN:AAA82237.1; PID:G904034

R:Lunec, J.

submitted to the EMBL Data Library, August 1995

A:Description: Multiple alternate spliced mdm2 transcripts with loss of p53 binding dom

A:Reference number: G09070

A:Accession: G02026

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-27, 223-491 <LUN>

A:Cross-references: EMBL:U33199; NID:G992676; PIDN:AAA75514.1; PID:G992677

A:Experimental source: splice form A

C:Genetics:

A:Gene: GDB:MDM2

A:Cross-references: GDB:250456; OMIM:164785

A:Map position: 12q14.3-12q15

C:Superfamily: human p53-binding protein mdm2

C:Keywords: alternative splicing; oncogene; phosphoprotein

P:1-491/Product: p53-binding protein mdm-2 #status predicted <MAT1>

F:1-27, 223-491/Product: p53-binding protein mdm-2, splice form A #status predicted <MAT

Query Match 99.0%; Score 1084.5; DB 1; Length 491;

Best Local Similarity 99.5%; Pred. No. 2.3e-80;

Matches 216; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Oy 1 MCNTNMSVPTDGAVTTSQIPASSEQETLVRPKLLKLLKSVGAQKDYTMKEVLVYL-QY 59

Db 1 MCNTNMSVPTDGAVTTSQIPASSEQETLVRPKLLKLLKSVGAQKDYTMKEVLVYLQY 60

QY 60 IMTKRLYDEKQOHHVYCSNDLGLFGVPSFVKHEHRIYTMIVRNLVVNVQSSDSGT 119
DB 61 IMTKRLYDEKQOHHVYCSNDLGLFGVPSFVKHEHRIYTMIVRNLVVNVQSSDSGT 120
QY 120 SVSENCHLEGSGDQKLVQELQBEKPSSSHLSVSRPSTSSRRRAISSETEENSDELSEGRQ 179
DB 121 SVSENCHLEGSGDQKLVQELQBEKPSSSHLSVSRPSTSSRRRAISSETEENSDELSEGRQ 180
QY 180 RKRHKSISLSFDESIALCVIREICERSSSESTG 216
DB 181 RKRHKSISLSFDESIALCVIREICERSSSESTG 217

RESULT 2
S15349
mdm2 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 31-Mar-2000
C:Accession: S15349
R:Fakhrazadeh, S.S.; Trusko, S.P.; George, D.L.
EMBO J. 10, 1565-1569, 1991
A:Title: Tumorigenic potential associated with enhanced expression of a gene that is and
A:Reference number: S15349; MUID:91224107; PMID:2026149
A:Accession: S15349
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-489 <PAK>
A:Cross-references: EMBL:X58876; NID:953038; PIDN:CAA41684.1; PID:G53039
C:Genetics:
A:Gene: mdm2
C:Superfamily: human p53-binding protein mdm2

Query Match 73.5%; Score 804.5; DB 2; Length 489;
Best Local Similarity 76.0%; Pred. No. 1.2e-57;
Matches 168; Conservative 16; Mismatches 24; Indels 13; Gaps 4;

QY 1 MCNTNMSVPTDGAATTSQIPASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLYL-QY 59
DB 1 MCNTNMSVTEGAATTSQIPASEQETLVPRPKLLKLLKSVGAQNDYTMKEIIFYGQY 60

QY 60 IMTKRLYDEKQOHHVYCSNDLGLFGVPSFVKHEHRIYTMIVRNLVVNVQSSDSGT 119
DB 61 IMTKRLYDEKQOHHVYCSNDLGLFGVPSFVKHEHRIYTMIVRNLVVNVQSSDSGT 117

QY 120 SVSENCHLEGSGDQKLVQELQBEKPSSSHLSVSRPSTSSRRRAISSETEENSDELSEGRQ 179
DB 118 SLSESRQPEGSDLKPLQAPPBEKPSDDLISRLSTSSRRRSISSETEENTDELPGERH 177

QY 180 RKRHKSISLSFDESIALCVIREICERSSSEST 215
DB 178 RKRHKSISLSFDESIALCVIREICERSSSEST 214

RESULT 3
S17195
MDM2-like p53-binding protein MDMX - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 31-Mar-2000
C:Accession: S17195
R:Shvarts, A.; Steegenga, W.T.; Riteco, N.; van Laar, T.; Dekker, P.; Bazuine, M.; van H
EMBO J. 15, 5349-5357, 1996
A:Title: MDMX: a novel p53-binding protein with some functional properties of MDM2.
A:Reference number: S17195; MUID:97050840; PMID:8895579
A:Accession: S17195
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-489 <SHV>
A:Cross-references: EMBL:AF007110; NID:G2253388; PIDN:AAB62927.1; PID:G2253389
C:Genetics:
A:Gene: MDMX
C:Function:
A:Description: inhibits transcription activation function of tumour suppressor protein p
C:Superfamily: human p53-binding protein mdm2

Query Match 25.3%; Score 277; DB 2; Length 489;
Best Local Similarity 36.9%; Pred. No. 7.4e-15;
Matches 79; Conservative 29; Mismatches 68; Indels 38; Gaps 7;

QY 21 ASEQETLVPRPKLLKLLKSVGAQKDYTMKEVLYL-QYIMTKRLYDEKQOHHVYCSND 79
DB 20 SSEQISQVRPKLQLLKILHAAGAQGEVFTMKEVLMHYLGQYIMVKQLYDQQQHMVYCGGD 79

QY 80 LLGDLFGVPSFVKHEHRIYTMIVRNLVVNVQSSDSGTSVSENCHLEGSGDQKLVQ 139
DB 80 LLGDLFGVPSFVKHEHRIYTMIVRNLVVNVQSSDSGTSVSENCHLEGSGDQKLVQ 124

QY 140 ELQBEKPSSSHLSVSRPSTSSRRRAISSETEENSDELSEGRQ---KRHKSISLSFDES 195
DB 125 DHTMDPSPQDLKKGATEYSNPKRTE-EEDTHLTPTSRHKCRDRADEDLIEHLSQDET 183

QY 196 LALCV-----IREICERSSSS 212
DB 184 SRLDLDFFBEMDVAGLPFWFLNLRNCPKNSG 217

RESULT 4
A59284
unconventional myosin heavy chain - Tetrahymena thermophila (fragment)
C:Species: Tetrahymena thermophila
C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 23-Mar-2001
C:Accession: A59284
R:Garces, J.; Gavin, R.H.
Journal of Eukaryot. Microbiol. 45, 252-259, 1998
A:Title: A PCR screen identifies a novel, unconventional myosin heavy chain gene (MYO1)
A:Reference number: A59284; MUID:98291456; PMID:9627986
A:Accession: A59284
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-851 <GAR>
A:Cross-references: GB:U87268; NID:G3258593; PIDN:AAC24454.1; PID:G3258594
A:Experimental source: gene MYO1
C:Genetics:
A:Gene: MYO1
A:Genetic code: SGC5
C:Superfamily: myosin motor domain homology
P:63-776/Domain: myosin motor domain homology <NMO>

Query Match 9.5%; Score 103.5; DB 2; Length 851;
Best Local Similarity 21.7%; Pred. No. 1.7;
Matches 54; Conservative 36; Mismatches 84; Indels 75; Gaps 11;

QY 5 NMSVPTDGAATTSQIPASEQETLVPRPKL---LLKLLKSVGAQ---KDYTMKEVLYLQY 59
DB 364 NMSLPF---HKEFKNASELLQVNPKEESLVFKRVKGVTTVIKSPQTAECLE-SMRD 418

QY 60 IMTKRLYDEKQOHHVYCSN-----DLIG-----DLFGVPSFVKHEHRI----- 98
DB 419 SLKKNLYSLFNLVWKLQNLLPQDLNLSLGLDIFGFSEFINSFQOLCINFTN 478

QY 99 ---YTMIV-----RNLVVNVQSSDSGTSTV-SENCHL 128
DB 479 EKFPQQIVQVYFKSEEDYKQGLDVFSDYQENQLIIDTIKKPNGIMVLLDENCSL 538

QY 129 EGSDQKLVQELQBEKPSSSHLSVSRPSTSSRRRAISSETEENSDELSEGRQ 176
DB 539 GSGTDDSFNLQITKNH--SKNQLVSPFKDKYSFSLSHTKRKYVYDVRGFRDKNDKDEVST 596

QY 177 ERQRKHKHS 185
DB 597 EVEKSLQSS 605

RESULT 5
T30336
nuclear/mitotic apparatus protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 21-Jul-2000
C/Accession: T30336
R/Werdes, A.; Ramyar, K.; Vechio, J.D.; Cleveland, D.W.
Cell 87, 447-458, 1996
A/Title: A complex of NuMA and cytoplasmic dynein is essential for mitotic spindle assembly
A/Reference number: Z20828; MUID:97053784; PMID:8898198
A/Accession: T30336
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-2253 <NR>
A/Cross-references: EMBL:X07624; NID:G1514670; PIDN:CAA68905.1; PID:G1514671
C/Genetics:
A/Gene: NuMA

Query Match 9.1%; Score 99.5; DB 2; Length 2253;
Best Local Similarity 24.3%; Pred. No. 12;
Matches 58; Conservative 37; Mismatches 99; Indels 45; Gaps 8;
QY 10 TDGAVTTSQIPASEQETLVRPKPLLLKLLKSVGAQKQVTYMKVLFYQYIMTKRLYDEK 69
Db 764 TAGKTESQLKHEEBYQXANESLQAKLAGSCAAIKQREBERDELKSVVDIW-KAKYGES 822
QY 70 QOHIVYCSNDLLGDLFGVPSFVKERKXYMYRNL-----VVNQESSDSGTSVSE 123
Db 823 QCKIAQNSCHM-----QEQTLELKTSHSDYQOEGERSKVLMEKASSETSSQLE 874
QY 124 -----NRCHLEGGSDQKDLVQELQ--EEKPSSSHLVSRPSTSSRRRAISETEE 169
Db 875 KINQLSGELSAANACIKERAEEKLVSAHSAEKKLKYAQGESERLSHLETALSNKQ 934
QY 170 NSP-----ELSGERQ-----KRHKSQSI-SLSPDESALCVREICERSSSS 212
Db 935 DLDCLAKELSDERYKAEFAEMVKVLKQNSERIASLESLXNLSLAVKVERKCESEKUS 993

RESULT 6

SYCHAL
5-aminolevulinate synthase (EC 2.3.1.37) precursor, nonspecific, mitochondrial - chicken
N/Alternate names: delta-aminolevulinate synthase, nonspecific
C/Species: Gallus gallus (chicken)
C/Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 05-May-2000
C/Accession: A23538; A23111; E31452
R/Naguire, D.J.; Day, A.R.; Borthwick, I.A.; Srivastava, G.; Wigley, P.L.; May, B.K.; El
Nucleic Acids Res. 14, 1379-1391, 1986
A/Title: Nucleotide sequence of the chicken 5-aminolevulinate synthase gene.
A/Reference number: A23538; MUID:86148479; PMID:3005973
A/Accession: A23538
A/Molecule type: DNA
A/Residues: 1-635 <WAG>
A/Cross-references: GB:X03517; NID:G63040; PIDN:CAA27223.1; PID:G763098
A/Experimental source: embryo liver
R/Borthwick, I.A.; Srivastava, G.; Day, A.R.; Pirola, B.A.; Snowell, M.A.; May, B.K.; E
Eur. J. Biochem. 150, 481-484, 1985
A/Title: Complete nucleotide sequence of hepatic 5-aminolevulinate synthase precursor.
A/Reference number: A23111; MUID:85257679; PMID:3839458
A/Accession: A23111
A/Molecule type: mRNA
A/Residues: 1-52, 'A', 54-635 <BOR>
A/Cross-references: GB:X02827; GB:M24366; NID:G63607; PIDN:CAA26595.1; PID:G63608
A/Experimental source: embryo liver
R/Riddle, R.D.; Yamamoto, M.; Engel, J.D.
Proc. Natl. Acad. Sci. U.S.A. 86, 792-796, 1989
A/Title: Expression of delta-aminolevulinate synthase in avian cells: separate genes en
A/Reference number: A31452; MUID:89128863; PMID:2915978
A/Accession: B31452
A/Status: preliminary; not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-52, 'A', 54-635 <RID>
A/Experimental source: liver
C/Comment: The mature functional enzyme is a dimer of identical chains containing pyrid
rate-limiting step in porphyrin biosynthesis.
C/Genetics:
A/Gene: ALASN

A/Introns: 67/1; 136/1; 188/1; 262/2; 324/1; 384/1; 439/1; 528/3; 583/1
C/Superfamily: 5-aminolevulinate synthase; glycine C-acyltransferase homology
C/Keywords: acyltransferase; coenzyme A; liver; mitochondrial matrix; mitochondrion; pt
F1-56/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F157-635/Product: 5-aminolevulinate synthase, nonspecific #status predicted <MAT>
F1243-583/Domain: glycine C-acyltransferase homology <GCA>
F1440/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 8.8%; Score 96.5; DB 1; Length 635;
Best Local Similarity 20.9%; Pred. No. 4.5;
Matches 65; Conservative 37; Mismatches 84; Indels 125; Gaps 14;
QY 8 VPTDGAVTTSQIPASEQETLVRP-KPLLLK-----LLKSVGA----- 43
Db 142 IPTNSVVRNTEAGEEGEQSLKKFKDMLKQRPESVSHLLQDNLPKSVTFQVDQPFPEKK 201
QY 44 ----QKD-----TYTMKEVLF-----YLQYIMTKRLYDEKQOHIVYCSNDLLG----- 82
Db 202 IDEKKKHDTYRVFTVNRKQKIFPMADDYSDSLITKK-----EVSVCNSDYLGMSRHP 255
QY 83 -----DLFGVPSFSV---KZ-----HRKIYTYMYRNLVVNQ--- 111
Db 256 RVCGAVMDTLKQHGAGAGGTNRISGTSKPHVDLEKELADLHGKDAALLFSFCFVANDSTL 315
QY 112 -----QESSDSGTSVS-----ENRC--HLEGGSDQKDLVQELQOEKRPSSSHLVS 153
Db 316 FTLAKMLPGCIIYSDSGNHASMIQGIINRSRVPRHIFRHDVNHURELLKSKDPSPTKIVA 375
QY 154 RPSTSSRRRAISETEENSDELS-----GERQKRHKSDSIS 189
Db 376 FETVHSDMGAVCPLEELCDVAHEGALTFFDEVHAGVLYGARGGGIGDRDGVHMKMDIIS 435
QY 190 LSPDESALCV 200
Db 436 GTLGKAFK-CV 445

RESULT 7

GB4559
probable NAM (no apical meristem)-like protein [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: GB4559
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: AB4420; MUID:20083487; PMID:10617197
A/Accession: GB4559
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-365 <STO>
A/Cross-references: GB:AB002093; NID:G406812; PIDN:AAD20120.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2G18060
A/Map position: 2
Query Match 8.5%; Score 93; DB 2; Length 365;
Best Local Similarity 20.6%; Pred. No. 4.4;
Matches 42; Conservative 39; Mismatches 81; Indels 42; Gaps 6;
QY 39 KSVGAQKDTYTMKEVLFY-----LQYIMTKRLYDEKQOHIVY-----CSNDLLGLDIFG 86
Db 159 RATGQAKNTEWSSSYFYDEVPAGVNSVMDPIDYISKQHNIFGKGLMKQLEFGMVDG 218
QY 87 VPSFSVKEHRKIYTYMYRNLVVNQESSDSGTSVSEN-----RCHLEGGSD 133
Db 219 INYIQSNQFIQLPQLQSPSLFMKRPSSMSITSDNNYNYKLPADSESPESFIRGDR 278
QY 134 QXDLVQEL-----QEEKPSSSHLVSR---PSTSSRRRAISETEENSEDE-----LSG 176
Db 279 RKKKQVMMTGNWRDLDFVASQLMSQEDNGTSSPAGHHIYNEDKNNNDVEMDSMFJSE 338

QY 177 ERQRKRKSDSISLFDLSALCV 200
| : : : : :
Db 339 REENRFVSEFLSTNSDYDIGCV 362
| : : : : :
RESULT 8
S35575
C:Species: Gallus gallus (chicken)
C:Date: 10-Dec-1993 #sequence revision 10-Nov-1995 #text_change 07-May-1999
C:Accession: S35575; S28931; S29162
R:Watanabe, B.
Biol. Chem. Hoppe-Seyler 374, 445-454, 1993
A:Title: Primary structure of subfragment-2 from adult chicken cardiac ventricular muscle
A:Reference number: S35575; MUID:94030661; PMID:8216895
A:Accession: S35575
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-520 <WAT>
R:Watanabe, B.
Biol. Chem. Hoppe-Seyler 373, 1045-1054, 1992
A:Title: Amino-acid sequence of the short subfragment-2 in adult chicken cardiac muscle
A:Reference number: S28931; MUID:93039740; PMID:1418675
A:Accession: S28931
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-259 <WAZ>
R:Watanabe, B.; Tanigawa, M.
Biol. Chem. Hoppe-Seyler 374, 27-35, 1993
A:Title: Primary structure of the hinge region in adult chicken cardiac myosin subfragment
A:Reference number: S29162; MUID:93176322; PMID:8439395
A:Accession: S29162
A:Molecule type: protein
A:Residues: 301-443 <WAZ>
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: ATP; cardiac muscle; heart
Query Match 8.4%; Score 92; DB 2; Length 520;
Best Local Similarity 21.3%; Pred. No. 8.1;
Matches 42; Conservative 40; Mismatches 63; Indels 52; Gaps 7;
QY 37 LKSVGAQKDTYTMKEVLFLQYIMTK-----RLYDERQKHIVYCSNDLLGDLFGVPSPS 91
| : : : : :
Db 5 LKSAETKEWANNKEEFLKLEAKSEARRKELEEQVSLVQEKNDLL----- 54
| : : : : :
QY 92 VKEHRIYTYMYRNLVVNQESSDGSSTSVSENCHL-----EGGSDQKDLVQELQEEK 145
| : : : : :
Db 55 -----LQQAQEDTLADAEERCDLLIKSKIQAAYKELTERVEDDE 96
| : : : : :
QY 146 PSSHLVSRPSTSSRR--AISTEENSDEL--GERQKRKSDSISLFDLSALC 199
| : : : : :
Db 97 ENVSEL-----TSKKKLEDECSLKKOIDDLEITLAKVEKEKATENKVNLTETMATL 151
| : : : : :
QY 200 --VIREICORSSSES 214
| : : : : :
Db 152 DENISKLTKEKSLQEA 168
| : : : : :
RESULT 9
G64594
hemolysin secretion protein precursor - Helicobacter pylori (strain 26695)
C:Species: Helicobacter pylori
C:Date: 09-Aug-1997 #sequence revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: G64594
R:Tombs, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenna, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Nature 368, 539-547, 1997
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: G64594

A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-433 <TOM>
A:Cross-references: GB:AE000573; GB:AE000511; NID:G2313712; PIDN:AAD07662.1; PID:G2313.
Query Match 8.4%; Score 91.5; DB 2; Length 433;
Best Local Similarity 24.3%; Pred. No. 7.1;
Matches 36; Conservative 31; Mismatches 64; Indels 17; Gaps 7;
QY 45 KDTYTMKEVLFLQYIMTKRLYD--EKQOHIVYCSNDLLGDLFGVPSPSVKEHR--KIVT 100
| : : : : :
Db 286 KNNMIVAQAQAKYTIYNNRNVFCGLAKLDHVVVF--KNNLYGMVFGNLNSFDITSHKSCRLGK 344
| : : : : :
QY 101 MIYRNLVVNQESSDGSSTSVSENCHLEGGSDQKDLVQELQEEKPSS--HLVSRPS 156
| : : : : :
Db 345 WYIEG--ACKENFNTSGYRALESH--HASVHAENDLVKAVQEDHITSKYLEHKVHLME 401
| : : : : :
QY 157 TSSRRRAISTEENSDELSEGERQKRHK 184
| : : : : :
Db 402 DSAK-----HVRENIDKMFYEKQDELNK 424
| : : : : :
RESULT 10
H71917
methyl-accepting chemotaxis protein (MCP) - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C:Date: 12-Feb-1999 #sequence revision 12-Feb-1999 #text_change 08-Oct-1999
C:Accession: H71917
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.; Ives, C.; Gibson, R.; Marberg, D.; Malls, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.; Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric pat
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: H71917
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <ARN>
A:Cross-references: GB:AE001487; GB:AE001439; NID:G4155086; PIDN:AAD06131.1; PID:G41551
A:Experimental source: strain J99
C:Genetics:
A:Gene: jhp0546
Query Match 8.3%; Score 90.5; DB 2; Length 433;
Best Local Similarity 23.6%; Pred. No. 8.6;
Matches 33; Conservative 33; Mismatches 63; Indels 17; Gaps 7;
QY 45 KDTYTMKEVLFLQYIMTKRLYD--EKQOHIVYCSNDLLGDLFGVPSPSVKEHR--KIVT 100
| : : : : :
Db 286 KNNMIVAQAQAKYTIYNNRNVFCGLAKLDHVVVF--KNNLYGMVFGNLNSFDITSHKSCRLGK 344
| : : : : :
QY 101 MIYRNLVVNQESSDGSSTSVSENCHLEGGSDQKDLVQELQEEKPSS--HLVSRPS 156
| : : : : :
Db 345 WYIEG--ACKENFNTSGYRALESH--HASVHAENDLVKAVQEDHITSKYLEHKVHLME 401
| : : : : :
QY 157 TSSRRRAISTEENSDELSEGERQKRHK 184
| : : : : :
Db 402 DSAK-----HVRENIDKMFYEKQDELNK 424
| : : : : :
RESULT 11
T16881
hypothetical protein T14G12.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16881
R:Wilcox, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid T14G12.
A:Reference number: Z18596
A:Accession: T16881
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA

A:Residues: 1-2037 <WIL>
A:Cross-references: EMBL:U41368; NID:g1086843; PID:g1086848; PIDN:AAA82437.1; CESP:T14G1
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
C:Accession: T01449
R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Con-
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, January 1998
A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.
A:Reference number: Z14211
A:Accession: T01449
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-751 <SHI>
A:Cross-references: EMBL:AC003113; NID:g2689438; PID:g2781355; GSPDB:GNO0059; ATSP:F2401
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:F2401.11
A:Map position: 1

Query Match 8.2%; Score 89.5; DB 2; Length 751;
Best Local Similarity 24.7%; Pred. No. 20; Mismatches 37; Indels 31; Gaps 9;
Matches 46; Conservative 37

Qy 24 QETLVPRPKLLKLLKSVCAQKDTYMKVLFYQYIMTKELYDEKQKHIVYCSND 79
Db 284 QEKIRWRP--LKPVVDHDIRLGQMPVNCRFKE---LREIVSSRFPSSKAVLIKYKND 337

Qy 80 LLGDLFGVPSPFVKHKKIYTMIRNLVVNQSSDSGTSVSNRCHLEGGSQKD--L 137
Db 338 --GLVITITSTA-----ELKLAESAADCILTKPEPTDKSDSVGMRLHVVVDVSPQEPL 390

Qy 138 VOELQ---EEKPSSHLVSRPSTSSRRRAISPTNSDELSSGERORKEKHS----DSISL 190
Db 391 LSEEEVEEKKVIEVISPTSES-----LSSTEINTKTKVEKEKASSDEPTEKEL 445

Qy 191 SFDESL 196
Db 446 EMDDWL 451

RESULT 13
T04227
hypothetical protein F14M19.10 - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cross)

A:Residues: 1-820 <BEV>
A:Cross-references: EMBL:AL049480
A:Experimental source: cultivar Columbia; BAC clone F14M19
C:Genetics:
A:Map position: 4
A:Introns: 158/3; 227/3; 243/2; 295/3; 316/2; 379/3; 409/3; 456/2; 495/3
A:Note: F14M19.10

Query Match 8.1%; Score 89; DB 2; Length 820;
Best Local Similarity 22.4%; Pred. No. 25; Mismatches 48; Conservative 34; Indels 54; Gaps 9;
Matches 48

Qy 23 EQETLVPRPKLLKLLKSVGAQKD-----TYTMKEVLFYQYIMTKELYDE 68
Db 328 EKKEVAKPEP-----DVGKDEENEDDKLLNELEELTNVDKTKKQAKKILAKRAKD 380

Qy 69 KQKHIVYCSNDLIGD-----LFGVPSPFVKHKKIYTMIRNLVVNQSSDSGTSV- 121
Db 381 KARKATGFMVDLEDFVDNELFSLNAIKGK-----KDLMAVDNDE--DNGNAVD 429

Qy 122 SENRCHLEGGS--PQKOLVQELQEEKPS-----SSHLVSRP-STSSRRRAISE 166
Db 430 SENEDHGEAGSDSDSDRSDDEERQKYTFQMEIEFEQAVRYVVKKEGSAKQKQARQA 489

Qy 167 TENSDLSGERORRKHKSISLSDFESLALCV 200
Db 490 HAEKLEEGDGEEMKIDYDSMNEEKDEANPLVV 523

RESULT 14
C71339
probable DNA mismatch repair protein (mutS) - syphilis spirochete
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C:Accession: C71339
R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwi-
rson, J.; Khataak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; Mcd
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A:Reference number: A71250; MUID:98332770; PMID:9665876
A:Accession: C71339
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-900 <COL>
A:Cross-references: GB:AE001212; GB:AE000520; NID:g3322597; PIDN:AA65315.1; PID:g33226
A:Experimental source: strain Nichols
C:Genetics:
A:Gene: TP0328
C:Superfamily: DNA mismatch repair protein mutS

Query Match 8.0%; Score 88; DB 2; Length 900;
Best Local Similarity 21.9%; Pred. No. 34; Mismatches 48; Conservative 36; Indels 56; Gaps 8;
Matches 48

Qy 31 KPILLKLLKS-----VGAQK-----DTYTMKEVLFYQYIMTKELYDEKQ-- 70
Db 336 RPHILKTRATLSCVRDVERLVARVALEKAHGDLLALKE---SLRAILFRSLERESPF 392

Qy 71 -QHIVYCSND--LLGDLFGVPSPFVKHKKIYTMIRNLVVNQSSDSGTSVSNRCH 127
Db 393 PPDLLPSEGDTFVLQELYGLLEQSIKEDCPV-TLSDGNLI-----KEGFSASLDELH 443

Qy 128 LEGSGDQKOLVQELQEEKP-----SSHLVSRPSTSSRRRAISE 166
Db 444 RVRDNEAILKQYLAEERERTGICTLKMYKRMGLHFLEYSKGLSAVFAHFRIRRLSLN 503

A:Residues: 1-2037 <WIL>
A:Cross-references: EMBL:U41368; NID:g1086843; PID:g1086848; PIDN:AAA82437.1; CESP:T14G1
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
C:Accession: T01449
R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Con-
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, January 1998
A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.
A:Reference number: Z14211
A:Accession: T01449
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-751 <SHI>
A:Cross-references: EMBL:AC003113; NID:g2689438; PID:g2781355; GSPDB:GNO0059; ATSP:F2401
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:F2401.11
A:Map position: 1

Query Match 8.2%; Score 89.5; DB 2; Length 751;
Best Local Similarity 24.7%; Pred. No. 20; Mismatches 37; Indels 31; Gaps 9;
Matches 46; Conservative 37

Qy 24 QETLVPRPKLLKLLKSVCAQKDTYMKVLFYQYIMTKELYDEKQKHIVYCSNDLIGD 83
Db 48 QPVLKRGSL-----KDVY-HGQVQHLARLOPKSLTAKDITMLERELLNH 95

Qy 84 LFGVPSPFVKHKKIYTMIRNLVVNQSSDSGTSVSNRCHLEGGSQD----- 134
Db 96 GFIITSDPQE-----LVLTPEYETSEDTSSVDVLLASFGAFDNLRETWG 144

Qy 135 KOLVQELQEEKPS-----SSHL-----VSRPSTSSRRR-----AISFTEE 169
Db 145 SDHVQ-IDHQNPNPRVTIPSAILSPPTNGSHLNVTVQSPSLTVHRDSDALGSLSRQPS 203

Qy 170 NSDELSSGERORR 182
Db 204 LADELQDERHQRR 216

RESULT 12
T01449
cytoskeletal protein homolog F2401.11 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 23-Mar-2001
C:Accession: T01449
R:Shinn, P.; Buehler, E.; Dewar, K.; Feng, J.; Kim, C.; Li, Y.; Sun, H.; Conway, A.; Con-
eologis, A.; Ecker, J.R.
submitted to the EMBL Data Library, January 1998
A:Description: Genomic sequence for Arabidopsis thaliana BAC F2401.
A:Reference number: Z14211
A:Accession: T01449
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-751 <SHI>
A:Cross-references: EMBL:AC003113; NID:g2689438; PID:g2781355; GSPDB:GNO0059; ATSP:F2401
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: ATSP:F2401.11
A:Map position: 1

Query Match 8.2%; Score 89.5; DB 2; Length 751;
Best Local Similarity 24.7%; Pred. No. 20; Mismatches 37; Indels 31; Gaps 9;
Matches 46; Conservative 37

Qy 24 QETLVPRPKLLKLLKSVCAQKDTYMKVLFYQYIMTKELYDEKQKHIVYCSND 79
Db 284 QEKIRWRP--LKPVVDHDIRLGQMPVNCRFKE---LREIVSSRFPSSKAVLIKYKND 337

Qy 80 LLGDLFGVPSPFVKHKKIYTMIRNLVVNQSSDSGTSVSNRCHLEGGSQKD--L 137
Db 338 --GLVITITSTA-----ELKLAESAADCILTKPEPTDKSDSVGMRLHVVVDVSPQEPL 390

Qy 138 VOELQ---EEKPSSHLVSRPSTSSRRRAISPTNSDELSSGERORKEKHS----DSISL 190
Db 391 LSEEEVEEKKVIEVISPTSES-----LSSTEINTKTKVEKEKASSDEPTEKEL 445

Qy 191 SFDESL 196
Db 446 EMDDWL 451

RESULT 13
T04227
hypothetical protein F14M19.10 - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cross)

Job time : 31.7931 secs

QY 167 TEENDELGGRQRKHKSDSISLSDSLALCVIREIC 205
DB 504 ADRTTQLSEAKLAREGLVSPFQELFADIRRTVC 542

RESULT 15

VEHULC
Lamin C - human
C:Species: Homo sapiens (man)
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 10-Dec-1999
C:Accession: A02962; A24249
R:McKeon, F.D.; Kirschner, M.W.; Caput, D.
Nature 319, 463-468, 1986
A:Title: Homologies in both primary and secondary structure between nuclear envelope and
A:Reference number: A02962; MUID:86118697; PMID:3453101
A:Accession: A02962
A:Molecule type: mRNA
A:Residues: 1-572 <MCK>
A:Cross-references: GB:X03445; NID:G34235; PIDN:CAA27174.1; PID:G34236
R:Fisher, D.Z.; Chaudhary, N.; Blobel, G.
Proc. Natl. Acad. Sci. U.S.A. 83, 6450-6454, 1986
A:Title: CDNA sequencing of nuclear lamins A and C reveals primary and secondary structure
A:Reference number: A94121; MUID:86313596; PMID:3462705
A:Accession: A24249
A:Molecule type: mRNA
A:Residues: 1-572 <FIS>
A:Cross-references: GB:M33451; NID:G186925; PIDN:AAA36164.1; PID:G307108
C:Comment: Lamins A (see PIR:VEHULC) and C are products of alternative splicing of the
C:Comment: The lamins (A, B, and C) contains several alpha-helical domains capable of
C:Comment: The association of lamins, dependent upon ionic interactions, is interrupted
rates with lamin dissociation; it does not reform until telophase, when the lamins are
C:Genetics:
A:Gene: GDB:LMNA; LMN1
A:Cross-references: GDB:132146; OMIM:150330
A:Map position: 1q21.2-1q21.3
C:Function:
A:Description: structural component of the nuclear lamina, a fibrous meshwork on the nuc
C:Superfamily: Cytoskeletal keratin
C:Keywords: alternative splicing; coiled coil; membrane protein; nuclear membrane; phosp
F:2-572/Product: lamin C #status predicted <MAT>
F:2-33/Domain: head <HED>
F:34-388/Domain: rod <ROD>
F:34-70/Region: coil 1A
F:81-218/Region: coil 1B
F:243-388/Region: coil 2
F:266/Region: heptad change of phase
F:330/Region: heptad change of phase
F:389-572/Domain: tail <END>
F:417-420/Region: nuclear location signal

Query Match 8.0%; Score 87.5; DB 1; Length 572;
Best Local Similarity 21.5%; Pred. No. 21;
Matches 44; Conservative 39; Mismatches 69; Indels 53; Gaps 7;

QY 31 KPLLKLLKSVCAQKDTYTKVEVLFYQYIMTKLYDEKQOH---IVYCSN----- 78
DB 191 KQLQDEMLRRYDAENRLQTMKEELDFQKNIIYSEELRETKRHETRLVEIDNGKQREFESR 240
QY 79 --DLIGDLFGVPSFKRHRYITMYRNVLVNVNQESSDSGTS-----VSENRCHL 128
DB 241 LADALQELRAHQEDQVGYKKELEKTY-SAKLDNARQSAERNLNVGAHEELQOQSRI 299
QY 129 EGGSDQ-----KDLVQELQEKPSHLSVRPSTSSRRRAISETEENSDE 173
DB 300 DLSAQLSQLOKQLAAEKRLDLEDSLAREDRTS-----RLLAEKER---E 344
QY 174 LSGERQKRKHKSDSISLSFDESAL 198
DB 345 NAEWRARMQQDEYQELLDIKLAL 369

Search completed: March 14, 2004, 21:48:45

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 14, 2004, 21:53:34 ; Search time 357.517 Seconds
(without alignments)
2566.620 Million cell updates/sec

Title: US-10-057-510-4

Perfect score: 1095

Sequence: 1 MCNTNNSVPTDGAVTTSQIP.....ALCVIRICCRSSSSBSTG 216

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US10057510/runat_09032004.162234.25965/app_query.fasta_1.654
-DB=N Geneseq 29Jan04 -QPMF=fastap -SUFFIX=std.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10057510 -CGN 1.1.468 -runat_09032004.162234.25965 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -XGAPOP=6
-XGAPEXT=7 -XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:

1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1084.5	99.0	652	3 AAA75042	Aaa75042 cDNA enco
2	1084.5	99.0	852	2 AAQ92515	Aaq92515 Human dou
3	1084.5	99.0	852	2 AAQ87261	Aaq87261 Human dou
4	1084.5	99.0	1476	2 AAT61637	Aat61637 Murine do
5	1084.5	99.0	1476	4 AAC84596	Aac84596 Human MDM
6	1084.5	99.0	1476	8 ADA02668	Ada02668 Human MDM
7	1084.5	99.0	1476	9 ADB72406	Adb72406 Human MDM
8	1084.5	99.0	2372	2 AAQ49891	Aaq49891 Human MDM

9	1084.5	99.0	2372	2 AAQ94589	Aaq94589 Human MDM
10	1084.5	99.0	2372	2 AAT45151	Aat45151 Human MDM
11	1084.5	99.0	2372	2 AAT66410	Aat66410 Human MDM
12	1084.5	99.0	2372	2 AAT62065	Aat62065 Human MDM
13	1084.5	99.0	2372	2 AAV20549	Aav20549 Human MDM
14	1084.5	99.0	2372	2 AAV28876	Aav28876 Human MDM
15	1084.5	99.0	2372	2 AAV04836	Aav04836 cDNA sequ
16	1084.5	99.0	2372	2 AAV03607	Aav03607 cDNA sequ
17	1084.5	99.0	2372	2 AAV03607	Aav03607 Human MDM
18	1084.5	99.0	2372	2 AAZ37471	Aaz37471 Human sar
19	1084.5	99.0	2372	2 AAQ35093	Aaq35093 Nucleotid
20	1084.5	99.0	2372	3 AAA29389	Aaa29389 MDM2 onco
21	1084.5	99.0	2372	4 AAF80625	Aaf80625 Human p-5
22	1084.5	99.0	2372	4 AAD07530	Aad07530 Human mdm
23	1084.5	99.0	2372	5 AAS29240	Aas29240 Human mdm
24	1084.5	99.0	2372	6 ABL62113	AbL62113 Colon ade
25	1084.5	99.0	2372	6 AAD42713	Aad42713 Human dou
26	1084.5	99.0	2372	6 ABV94130	Abv94130 Breast ca
27	1084.5	99.0	2372	6 AAL43913	Aal43913 Human Dm2
28	1084.5	99.0	2372	8 ADA02667	Ada02667 Human MDM
29	1084.5	99.0	2372	9 ADB72405	Adb72405 Human MDM
30	1084.5	99.0	2372	9 ADD21436	Add21436 Human mdm
31	1084.5	99.0	2372	4 AAK51460	Aak51460 Human pol
32	833	76.1	2393	5 AAS94144	Aas94144 DNA encod
33	808	73.8	681	2 AAQ92516	Aaq92516 Human dou
34	808	73.8	681	2 AAQ87262	Aaq87262 Human dou
35	808	73.8	1302	2 AAQ92517	Aaq92517 Human dou
36	805	73.5	1302	2 AAQ87263	Aaq87263 Human dou
37	805	73.5	1470	9 ADD21759	Add21759 Mouse mdm
38	804.5	73.5	1470	8 ADA02665	Ada02665 Mouse mdm
39	804.5	73.5	1470	9 ADB72403	Adb72403 Mouse Mdm
40	804.5	73.5	1710	2 AAQ49892	Aaq49892 Murine MD
41	804.5	73.5	1710	2 AAQ94590	Aaq94590 Mouse MDM
42	804.5	73.5	1710	2 AAT45152	Aat45152 Murine MD
43	804.5	73.5	1710	2 AAT66411	Aat66411 Murine MD
44	804.5	73.5	1710	2 AAV20550	Aav20550 Mouse MDM
45	804.5	73.5	1710	2 AAV28877	Aav28877 Mouse MDM

ALIGNMENTS

RESULT 1

AAA75042

ID AAA75042 standard; cDNA; 652 BP.

XX AAA75042;

AC AAA75042;

XX 02-JAN-2001 (first entry)

DT 02-JAN-2001 (first entry)

DE cDNA encoding a human MDMIP-binding MDM2 polypeptide fragment.

XX Human; MDM2 interacting polypeptide; MDMIP; MDM2; cell cycle progression;
XX cell differentiation; cancer; glioma; squamous cell carcinoma;
XX breast cancer; astrocytoma; leukemia; lymphoma; tumorigenesis;
XX Gene therapy; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

PH CDS 1..652

FT /ttg= a

FT /transl_except= (pos: 172..177, aa: Gln)

FT /note= "partial sequence"

XX WO2000050590-A1.

XX 31-AUG-2000.

XX 23-FEB-2000; 2000WO-US004592.

XX 23-FEB-1999; 99US-0121192P.

XX 03-MAR-1999; 99US-0122643P.

XX 22-FEB-2000; 2000US-00510252.

XX (CURA-) CURAGEN CORP.
XX Nandabalan K, Yang M, Schulz VP;
XX WPI; 2000-558398/51.
XX P-PSDB; AAB08846.
XX Novel MDM2 interacting protein useful for treating or preventing
PT disorders involving aberrant levels of MDM2 and/or MDM-interacting
PT proteins, comprises a specific amino acid sequence.
XX Disclosure; Fig 2A; 78pp; English.
XX The present sequence encodes a fragment of a human MDM2 polypeptide,
CC which binds to a human MDM2 interacting polypeptide (MDMIP). The protein
CC fragment was used as bait in a yeast two hybrid system to identify MDMIP.
CC The MDMIP polypeptide is useful for detecting and removing MDM2
CC polypeptides in a biological sample by forming MDM2-MDMIP complexes.
CC MDMIP and MDM2 are useful to identify compounds or other agents which
CC modulate the activity of MDM2 and/or MDMIP-mediated processes. Agents
CC that modulate the function of MDMIP/MDM2 complexes are useful for
CC treating and preventing a disease or disorder involving aberrant levels
CC of MDM2 or MDMIP. MDMIP is also useful for treating diseases caused by
CC aberrant levels of expression of MDM2 genes, such as disorders of cell
CC cycle progression, cell differentiation, and transcriptional control,
CC including cancers such as human sarcoma, glioma, squamous cell carcinoma,
CC breast cancer, astrocytoma, leukemia and lymphoma, and tumorigenesis.
CC MDMIP and MDM2 nucleic acids are useful in gene therapy
XX
SQ Sequence 652 BP; 212 A; 116 C; 145 G; 179 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4.36e-116 Length: 652
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 3 Gaps: 1
US-10-057-510-4 (1-216) x AAA75042 (1-652)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrSerGlnIlePro 20
DB 1 ATGTGCAATACCAACATCTGTACTTACTGATGCTGTGCTGTAACACCTCACAGATTCCA 60
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLysLeuLeuLysSer 40
DB 61 GCTTCGGAAACAGACACCTGCTGTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120
QY 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 121 GTTGTGTGACAAAAGACACTTACTATGTAAGAGAGGTTCTTTTATCTTGGCCAGTAT 180
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 161 ATTATGACTTAACGATTATATGATGAGAGCAACACATATGTTATTTGTTCAATGAT 240
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 241 CTTCTAGGAGATTCTTGGCGTGCCAAAGCTTCTCTGTGAAAGAGACACAGGAAAATATAT 300
QY 100 ThrMetIleTyrArgAsnLeuValValLysGlnGlnGlnSerSerAspSerGlyThr 119
DB 301 ACCATGATCTACAGGAACCTTGTAGTAGTCAATACAGCAGGATCATCGACTCAGGTACA 360
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
DB 361 TCTGTGAGTGACACAGGGTGTCACTTGAAGTGGAGTGTATCAAAAGGACCTTGACAA 420
QY 140 GluLeuGlnGluLysProSerSerSerHisLeuValSerArgProSerThrSerSer 159
DB 421 GAGCTTCAGGAGAGAAACCTTCATCTTCACATTGGTTTCTAGACCACTACCTCACTCT 480

QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
DB 481 AGAAGGAGAGCAATTAGTGACACAGAGAAAATTCAGATGAATATCTGGTGAACGACAA 540
QY 180 ArgLysArgHisLysSerAspSerLysSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 541 AGAAACCCCAAAATCTGATAGTATTTCCCTTCCCTTGTATGAAGCCCTGGCTGTGT 600
QY 200 VallIleArgGluIleCysCysGluArgSerSerSerGluSerThrGly 216
DB 601 GTATTAAGGAGATATGTTGTGAAAGACAGTACGACGTGAATCTACAGGG 651
RESULT 2
AAQ92515
ID AAQ92515 standard; DNA; 852 BP.
XX
AC AAQ92515;
XX
DT 02-FEB-1996 (first entry)
XX
DE Human double minute gene 2 (hdm-2) fragment 1.
XX
KW Human double minute gene 2; hdm-2; antibody binding region; antigen;
KW cancer; sarcoma; rhabdomyosarcoma; diagnosis; immunoassay; ss.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..852
FT /tag= a
FT /label= Fragment 1
FT /note= "encodes amino acids 1-284 of hdm-2, i.e. the 5'-
FT region of the ORF only"
FT
XX
PN DE4339533-Al.
XX
PD 14-JUN-1995.
XX
PF 19-NOV-1993; 93DE-04339533.
XX
PR 19-NOV-1993; 93DE-04339533.
XX
(DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.
XX
PI Zentgraf H, Klein R, Frey M, Martens R;
XX
XX WPI; 1995-216248/29.
DR P-PSDB; AAR75494.
XX
PT Detection of human double minute gene 2 (hdm-2) antibodies - by
PT incubation with new hdm-2 or antibody-binding hdm-2 fragments; useful in
PT the detection of specific cancers.
XX
PS Claim 13; Fig 2; 12pp; German.
XX
CC DNA fragments coding for amino acids 1-284, 58-284 and 58-491 of the hdm-
CC 2 (human double minute 2) gene product are claimed. The overlapping
CC protein fragments contain binding regions for hdm-2-specific antibodies
CC and are useful for identifying such antibodies using a claimed
CC immunoassay method. The presence of anti-hdm-2 antibodies is diagnostic
CC of certain forms of cancer, e.g. rhabdomyosarcoma
XX
SQ Sequence 852 BP; 271 A; 145 C; 199 G; 237 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 6.48e-116 Length: 852
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 2 Gaps: 1

US-10-057-510-4 (1-216) x AAQ92515 (1-852)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
 Db 1 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTGTAACACCTCACAGATTCCA 60
 QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLysSer 40
 Db 61 GCTTCGGAACAAGAGACCCCTGGTGTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120
 QY 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu---GlnTyr 59
 Db 121 GTTGTGTCACAAAGACACTATATCTATGAAGAGGTCTCTTTTATCTTGGCCAGTAT 180
 QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyCysSerAsnAsp 79
 Db 181 ATTATGACTAAACGATTATATGATGAGAAGCAACACATATTGTATATTGTTCAATGAT 240
 QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
 Db 241 CTTCCTAGGAGATTGTTTGGCTGCCAGCTTCTCTGTGAAGAGACAGGAAATATAT 300
 QY 100 ThrMetIleTyArgAsnLeuValValAsnGlnGlnLysSerAspSerGlyThr 119
 Db 301 ACCATGATCTACAGGAACCTTGTGTAGTGTCAATCAGCAGGAATATCGACTCAGGTACA 360
 QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
 Db 361 TCTGTGAGTGAGAACAGGTGTCCACTTGAAGTGGGAGTGATCAAAAGGACCTTGTACAA 420
 QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
 Db 421 GAGCTTCAGAGAGATTGTTTGGCTGCCAGCTTCTCTGTGAAGAGACAGGAAATATAT 180
 QY 160 ArgArgAlaIleSerGluThrGluAsnSerAspGluLeuSerGlyGluArgGln 179
 Db 481 AGAAGGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATATCTGGTGAACGACAA 540
 QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
 Db 541 AGAAACGCCCAAAATCTCATAGTATTTCCCTTTCCTTTGATGAAGGCTGGCTCTGTGT 600
 QY 200 ValIleArgGluLeuCysCysGluArgSerSerSerSerSerSerGlyThrGly 216
 Db 601 GTAATAAGGGAGATATGTTGTGAAGAACAGTACAGCATGATCTACAGGG 651

RESULT 3

AAQ87261

ID AAQ87261 standard; DNA; 852 BP.

XX AC AAQ87261;

XX DT 25-MAR-2003 (revised)

XX DT 25-JAN-1996 (first entry)

XX DE Human double minute gene 2 (hdm-2) fragment 1.

XX DE Human double minute gene 2; hdm-2; antibody binding region; antigen;

XX KW cancer; sarcoma; rhabdomyosarcoma; diagnosis; ss.

XX OS Homo sapiens.

XX XX Key Location/Qualifiers

XX FT "1..852

XX FT /tag= a

XX FT /label= Fragment_1

XX FT /notes= "encodes amino acids 1-284 of hdm-2, i.e. the 5'-

XX FT region of the ORF only"

XX PN DE4345249-A1.

XX PD 24-MAY-1995.

XX XX

PF 19-NOV-1993; 93DE-04345249.

XX PR 19-NOV-1993; 93DE-04339533.

XX PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.

XX Zentgraf H, Klein R, Frey M, Martens R;

XX WPI; 1995-195167/26.

DR P-PSDB; AAR75397.

XX New hdm-2 fragments contg. antibody binding region - used to detect

XX specific antibodies for diagnosis of cancers, also new DNA sequences

XX encoding them.

XX Claim 4; Fig 2; 11pp; German.

XX DNA fragments coding for amino acids 1-284, 58-284 and 58-491 of the hdm-

XX 2 (human double minute 2) gene product are claimed. The overlapping

XX protein fragments contain binding regions for hdm-2-specific antibodies

XX and are useful for identifying such antibodies. The presence of anti-hdm-

XX 2 antibodies is diagnostic of certain forms of cancer, e.g.

XX rhabdomyosarcoma. (Updated on 25-MAR-2003 to correct PF field.)

XX SQ Sequence 852 BP; 271 A; 145 C; 199 G; 237 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6.48e-116 Length: 852

Score: 1084.50 Matches: 216

Percent Similarity: 99.54% Conservative: 0

Best Local Similarity: 99.54% Mismatches: 0

Query Match: 99.04% Indels: 1

DB: 2 Gaps: 1

US-10-057-510-4 (1-216) x AAQ87261 (1-852)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20

Db 1 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTGTAACACCTCACAGATTCCA 60

QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLysSer 40

Db 61 GCTTCGGAACAAGAGACCCCTGGTGTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120

QY 41 ValGlyAlaGlnLysAspThrTyThrMetLysGluValLeuPheTyLeu---GlnTyr 59

Db 121 GTTGTGTCACAAAGACACTTATCTATGAAGAGGTCTCTTTTATCTTGGCCAGTAT 180

QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyCysSerAsnAsp 79

Db 181 ATTATGACTAAACGATTATATGATGAGAAGCAACACATATTGTATATTGTTCAATGAT 240

QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99

Db 241 CTTCCTAGGAGATTGTTTGGCTGCCAGCTTCTCTGTGAAGAGACAGGAAATATAT 300

QY 100 ThrMetIleTyArgAsnLeuValValAsnGlnGlnLysSerAspSerGlyThr 119

Db 301 ACCATGATCTACAGGAACCTTGTGTAGTGTCAATCAGCAGGAATATCGACTCAGGTACA 360

QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139

Db 361 TCTGTGAGTGAGAACAGGTGTCCACTTGAAGTGGGAGTGATCAAAAGGACCTTGTACAA 420

QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159

Db 421 GAGCTTCAGAGAGATTGTTTGGCTGCCAGCTTCTCTGTGAAGAGACAGGAAATATAT 180

QY 160 ArgArgAlaIleSerGluThrGluAsnSerAspGluLeuSerGlyGluArgGln 179

Db 481 AGAAGGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATATCTGGTGAACGACAA 540

QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199

Db 541 AGAAACGCCAACATCTAGTATTTCCTTTCTGATGAAGCCTGGCTGTGT 600
 Qy 200 ValileArgGluileCysGluArgSerSerSerGluSerThrGly 216
 Db 601 GTAATAGGAGAGATATGTTGTGAAGAGAGCAGTAGCAGTGAATCTACAGG 651

RESULT 4
 AAT61637 standard; cDNA; 1476 BP.
 XX AAT61637;
 AC AAT61637;
 XX 16-JAN-1998 (first entry)
 DT Murine double minute 2 coding sequence.
 DE Mouse; Mdm2; murine double minute; phosphoprotein; binding; modulation;
 KW tumour suppressor; p53; oncogene; cell cycle arrest; p107; antagonist;
 KW inhibition; transcription factor; adenocarcinoma; colon; cancer; breast;
 KW lung; stomach; myeloid leukaemia; lymphoma; hyperproliferative;
 KW restenosis; ss.
 XX Mus musculus.
 OS
 XX W09709343-A2.
 PN 13-MAR-1997.
 XX 02-SEP-1996; 96WO-FR001340.
 XX 04-SEP-1995; 95FR-00010331.
 XX (RHON) RHONE POULENC RORER SA.
 PA (INRM) INST NAT SANTE & RECH MEDICALE.
 XX Tocque B, Dubs-Poterszman M, Wasylyk B;
 XX WPT; 1997-192837/17.
 DR P-PSDB; AAW13600.
 XX Treating cancer with antagonist of oncogenic activity of protein Mdm2 -
 PT or nucleic acid encoding an antagonist, also viral vectors contg. this
 PT nucleic acid.
 XX Claim 2; Page 26-30; 43pp; French.

This is the nucleotide sequence encoding the mouse Mdm2 (murine double minute-2) protein, a 90 kD phosphoprotein which binds and modulates the activity of the tumour suppressor protein p53. It has now been shown that the mdm2 protein itself has oncogenic properties, especially in a p53-null background. Mdm2 is observed to unblock cell cycle arrest in G1 caused by over-expression of the p107 protein. This is especially done by the region covering amino acid 1-134. The invention therefore relates to antagonists able to inhibit the oncogenic activity of mdm2. These include fragments of the p53 protein, especially amino acids 1-52, 1-41, 6-41, 16-25 or 18-23 (AAW13602-6), or fragments of transcription factors e.g. TFIID, TBP or TAF250, which bind amino acids 1-134 of mdm2. Other inhibitors include compounds which disrupt binding to region 135-491 of mdm2, e.g. Rb, L5 or the transcription factor E2F. The antagonists are used to treat e.g. adenocarcinoma of the colon; cancer of the breast, lung or stomach; myeloid leukaemia; B cell lymphoma, or other hyperproliferative conditions such as restenosis

SQ Sequence 1476 BP; 491 A; 259 C; 332 G; 394 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 1.47e-115 Length: 1476
 Score: 1084.50 Matches: 216
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 0
 Query Match: 99.04% Indels: 1
 DB: 2 Gaps: 1

US-10-057-510-4 (1-216) x AAT61637 (1-1476)
 Qy 1 MetCysAenThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
 Db 1 ATGTGCATATACCAACATGTCTGTACCTACTCATGTGTCTGTAAACACCTCACAGATCCA 60
 Qy 21 AlaSerGluGlnGlnThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuLeuSer 40
 Db 61 GCTTCGGAAACAAGAGAGACCTGGTTAGACCAAGCCATTGCTTTTGAAGTATTATAAGTCT 120
 Qy 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
 Db 121 GTTGTGCACAAAAGACACTTATCTATGAAGAGGTTCTTTTATCTTGGCCAGTAT 180
 Qy 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
 Db 181 ATTATGACTAAACGATTATATGATGAGAAGCAACAACATATGTATATTTGTTCAATGAT 240
 Qy 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
 Db 241 CTTCTAGGAGATTGTTTGGCGTGCACAGCTTCTCTGTGAAGAGAGCAGAGAAATATAT 300
 Qy 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGlnSerSerAspSerGlyThr 119
 Db 301 ACCATGATCTACAGGAACCTTGGTAGTCAATCAGCAGGAATCATCGACTCAGGTACA 360
 Qy 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
 Db 361 TCTGTGAGTGAGAACAGGTGTCACTTGAAGGTGGAGTGATCAAAAGAGACCTTGTACAA 420
 Qy 140 GluLeuGlnGluGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
 Db 421 GAGCTTCAGGAAGAGAAACCTTCATCTTCACATTTGGTTCTAGACCATCTACCTCATCT 480
 Qy 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
 Db 481 AGARGGAGAGCAATTTAGTGAGACAGAGAGAAATTCAGATGAATATCTGTGTAACGACA 540
 Qy 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
 Db 541 AGAAACCCCAACAAATCTGATAGTATTTCCCTTTTGTATGAAGCCTGGCTCTGTGT 600
 Qy 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
 Db 601 GTAATAGGAGAGATATGTTGTGAAGAGAGCAGTAGCAGTGAATCTACAGG 651

RESULT 5
 AAC84596
 ID AAC84596 standard; DNA; 1476 BP.
 XX AAC84596;
 AC AAC84596;
 XX 02-APR-2001 (first entry)
 DT Human MDM2 protein encoding DNA.
 DE S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
 KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
 KW Bad; Bcl-2; tumour; cytosolic; ds.
 XX Homo sapiens.
 OS
 XX WO200075184-A1.
 PN 14-DEC-2000.
 PD 05-JUN-2000; 2000WO-US015449.
 XX 04-JUN-1999; 99US-0137494P.
 PR (UYVA) UNIV YALE.
 XX

PI Zhang H, Tsvetkov LM, Kondo T;
XX WPI: 2001-061703/07.
DR P-PSDB; AAB48284.
XX
XX Modulating polypeptide levels in a cell, diagnosing and treating tumor,
PT involves altering levels of proteins such as S-phase kinase associated
PT proteins 1, 2 and cullin/CDC53 proteins.
XX
XX Example; Page 93-95; 162pp; English.
XX
XX The invention relates to methods of altering the polypeptide levels in a
CC cell, using proteins selected from S-phase kinase associated proteins 1
CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
CC cullin/CDC53 family of proteins). The method is useful for altering the
CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
CC polypeptide in a cell. SKP2 and c-Myc-like protein levels are useful for
CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents
CC that modulate interactions between SKP and target proteins are useful for
CC treating tumours
XX
XX Sequence 1476 BP; 491 A; 259 C; 332 G; 394 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 1,476-115 Length: 1476
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 4 Gaps: 1

US-10-057-510-4 (1-216) x AAC84596 (1-1476)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrSerGlnIlePro 20
DB 1 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTAAACACCTCACAGATTCCA 60

QY 21 AlaSerGluClnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuSer 40
DB 61 GCTTCGGAACAAGACACCTGTTGTAGACCAAGACCATGCTTTTGAAGTTATTAAAGTCT 120

QY 41 ValGlyAlaGlnLysAspThrTyrrMetLysGluValLeuPheTyrrLeu---GlnTyrr 59
DB 121 GTTGTGTCACAAAAGACACTTACTATGAAGAGGTTCTTTTATCTGGCCAGTAT 180

QY 60 IleMetThrLysArgLeuTyrrAspGluLysGlnGlnHisLeuValTyrrCysSerAsnAsp 79
DB 181 ATTATGACTAAACGATTATATGATGAGAACCAACACATATTGTATATTGTTCAAATGAT 240

QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyrr 99
DB 241 CTTCTAGGAGATTGTTTGGCGTCCCAAGCTTCTCTGTGAAAGACACAGGAAATATAT 300

QY 100 ThrMetIleTyrrArgAsnLeuValValAlaGlnGlnGlnSerSerAspSerGlyThr 119
DB 301 ACCATGATCTACAGGAACCTGGTAGTCAATCAGCAGGATCATCGGACTCAGGTACA 360

QY 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139
DB 361 TCTGTGAGTGAGAACAGGTGTCCACCTTGAAGTGGGAGTGAACAAAGACACCTTGTACA 420

QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
DB 421 GAGCTTCAGGAAGAGAAACCTTCATCTTCATTTGGTTTCTAGACCATCTACCTCACT 480

QY 160 ArgArgArgAlaIleSerGluThrGluGluAenSerAspGluLeuSerGlyGluArgGln 179
DB 481 AGAAGGAGAGCAATTAAGTCAGACAGAGAAATTCAGATGAATTATCTGTTGAACGACAA 540

QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 541 AGAAACGCCACAAATCTCATAGTATTTCCCTTTCCTTTGATGAAAGCCTGGCTCTGTGT 600

QY 200 ValIleArgGluIleCysCysGluArgSerSerSerGluSerThrGly 216
DB 601 GTAATAAGGAGATATGTTGTGAAAGAGCAGTAGCAGTGAATCTACAGGG 651

RESULT 6
ADA02668
ID ADA02668 standard; cDNA; 1476 BP.
XX
XX ADA02668;
XX
XX 06-NOV-2003 (first entry)
XX
XX Human MDM2 carcinoma associated coding sequence, SEQ ID NO:1186.
XX
XX Human; carcinoma associated; oncogene; carcinoma; cancer; breast;
XX prostate; lymphoma; leukaemia; cytostatic; gene therapy; drug screening;
XX gene; ss.
XX
XX Homo sapiens.
XX
XX WO2003057146-A2.
XX
XX 17-JUL-2003.
XX
XX 26-DEC-2002; 2002WO-US041414.
XX
XX 26-DEC-2001; 2001US-00035832.
XX
XX (SAGR-) SAGRES DISCOVERY.
XX
XX Morris DW;
XX
XX WPI: 2003-587068/55.
XX
XX New recombinant nucleic acid encoding carcinoma associated protein,
XX useful for preparing compositions for treating carcinomas.
XX
XX Claim 1; SEQ ID NO 1186; 245pp; English.
XX
XX The invention relates to recombinant carcinoma associated (CA) nucleic
XX acid sequences from mouse and human (ADA01482-ADA03094), and to
XX recombinant carcinoma associated proteins (CAP) encoded by them. The
XX invention also encompasses expression vectors and host cells comprising a
XX CA nucleic acid, a polypeptide (especially an antibody) that specifically
XX binds to the protein, and a biochip comprising CA nucleic acid or
XX fragments thereof. The sequences of the invention were identified using
XX oncogenic retroviruses, which insert into the genome of the host organism
XX at random. Many of these do not carry transduced host oncogenes or
XX pathogenic trans-acting viral genes, meaning that cancer incidence is a
XX direct consequence of the effects of proviral integration into host
XX protooncogenes. The CA nucleic acid sequences can be used to diagnose
XX carcinoma (especially breast cancer, prostate cancer, lymphoma or
XX leukaemia) or a propensity to carcinoma by determination of the sequence
XX of a CA gene, or by determination of CA gene expression in particular
XX tissues. CA nucleic acids, proteins and antibodies are also useful as
XX therapeutic agents and in screening and evaluating drug candidates. The
XX present sequence represents a specifically claimed human CA nucleic acid
XX sequence of the invention. Note: The complete sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1476 BP; 491 A; 259 C; 332 G; 394 T; 0 U; 0 Other;
SQ

Alignment Scores:
Pred. No.: 1,476-115 Length: 1476
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 8 Gaps: 1

US-10-057-510-4 (1-216) x ADA02668 (1-1476)


```
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 1 ATGTGCATACCAACATGCTGTACCTACTGATGCTGTACCTACCTACAGATTCCA 60

QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysSer 40
DB 61 GCTTCGGAACAAGAGACCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120

QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 121 GTTGGTGACAAAAGACACTTATCTATGAAAGAGGTTCTTTTATCTTGCCAGTAT 180

QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 181 ATTATGACTAAACGATTATATGATGAGAACCAACATATTTATATTGTTCAAAATGAT 240

QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 241 CTTCTAGGAGATTGTTTGGCGTCCCAAGCTTCTCTGAAAGAGCACAGGAAATATAT 300

QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
DB 301 ACCATGATCTACAGAACTTGGTAGTATGATCAATCAGCAGGATCATCGGACTCAGGTACA 360

QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
DB 361 TCTGTGAGTGAGAACAGGTGTCACTTGAAGGTGGAGTGATCAAAAGGACCTTGTACAA 420

QY 140 GluLeuGlnGluLysProSerSerHisIleValSerValSerArgProSerThrSerSer 159
DB 421 GAGCTTCAGGAAGAGAACCTTCATCTTCACATTTGGTTTCTAGACCATCTACTCATCT 480

QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
DB 481 AGAAGGAGAGCAATTAGTGACACAGAGAAATTCAGATGAATTTATCTGTGTAACGACAA 540

QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 541 AGAAACGCCCAAAATCTGATGATTTCCCTTTTCTTGTATGAAGCCCTGGCTGTGTGT 600

QY 200 ValIleArgGluIleCysGluArgSerSerSerSerGluSerThrGly 216
DB 601 GTAATAGGGAGATATGTTGTGAAGAGAGCAGTAGCAGTGAATCTACAGGG 651
```

RESULT 7

ADB72406
ID ADB72406 standard; cDNA; 1476 BP.

AC ADB72406;

DT 04-DEC-2003 (first entry)

XX Human MDM2 cDNA.

DE human; ss: cytotostatic; gene therapy; vaccine; carcinoma; lymphomas;
KW cancer; neoplasm; adenocarcinoma; sarcoma; gene.

XX Homo sapiens.

OS WO2003008583-A2.

PN 30-JAN-2003.

XX 26-DEC-2001; 2001WO-US051291.

PF 02-MAR-2001; 2001US-00798586.

XX 23-OCT-2001; 2001US-00004113.

PR 08-NOV-2001; 2001US-00052482.

PR 30-NOV-2001; 2001US-00397722.

PR 20-DEC-2001; 2001US-00034650.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW, Engelhard EK;
XX WPI; 2003-239337/23.
XX New recombinant nucleic acid, useful for treating carcinomas, lymphomas,
XX cancers, neoplasm, adenocarcinoma, or sarcomas.
XX Claim 1; SEQ ID NO 234; 2304pp; English.

XX The invention relates to a novel recombinant nucleic acid comprising a
XX nucleotide sequence selected from any of the 660 sequences fully defined
XX in the specification. A polynucleotide of the invention has cytostatic
XX activity, and may have a use in gene therapy, or in a vaccine. The
XX recombinant nucleic acids and polypeptides are useful for treating
XX carcinomas, e.g. lymphomas, cancers, neoplasm, adenocarcinoma, and
XX sarcomas. The present sequence represents a human cDNA of the invention.

XX Sequence 1476 BP; 491 A; 259 C; 332 G; 394 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.47e-115 Length: 1476
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 9 Gaps: 1

US-10-057-510-4 (1-216) x ADB72406 (1-1476)

```
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 1 ATGTGCATACCAACATGCTGTACCTACTGATGCTGTACCTACCTACAGATTCCA 60

QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLysSer 40
DB 61 GCTTCGGAACAAGAGACCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 120

QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 121 GTTGGTGACAAAAGACACTTATCTATGAAAGAGGTTCTTTTATCTTGCCAGTAT 180

QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 181 ATTATGACTAAACGATTATATGATGAGAACCAACATATTTATATTGTTCAAAATGAT 240

QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 241 CTTCTAGGAGATTGTTTGGCGTCCCAAGCTTCTCTGAAAGAGCACAGGAAATATAT 300

QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
DB 301 ACCATGATCTACAGAACTTGGTAGTATGATCAATCAGCAGGATCATCGGACTCAGGTACA 360

QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
DB 361 TCTGTGAGTGAGAACAGGTGTCACTTGAAGGTGGAGTGATCAAAAGGACCTTGTACAA 420

QY 140 GluLeuGlnGluLysProSerSerHisIleValSerValSerArgProSerThrSerSer 159
DB 421 GAGCTTCAGGAAGAGAACCTTCATCTTCACATTTGGTTTCTAGACCATCTACTCATCT 480

QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
DB 481 AGAAGGAGAGCAATTAGTGACACAGAGAAATTCAGATGAATTTATCTGTGTAACGACAA 540

QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 541 AGAAACGCCCAAAATCTGATGATTTCCCTTTTCTTGTATGAAGCCCTGGCTGTGTGT 600

QY 200 ValIleArgGluIleCysGluArgSerSerSerSerGluSerThrGly 216
DB 601 GTAATAGGGAGATATGTTGTGAAGAGAGCAGTAGCAGTGAATCTACAGGG 651
```



```

DR P-PSDB; AAR76696.
XX New human MDM2 cDNA - used to develop prods. for use in the diagnosis and
PT treatment of tumours.
XX
XX Claim 1; Col 19-24; 34pp; English.
XX
XX The human MDM2 gene is genetically altered (i.e. amplified) in human
CC tumour cells. Detecting that the gene has become amplified or detecting
CC increased gene product expression (using probes, proteins, antibodies and
CC inhibitors) allows diagnosis and therapy of cancers such as colorectal
CC carcinoma, lung cancer and chronic myelogenous leukaemia. The human MDM2
CC protein binds to human p53 and allows the cell to escape from p53-
CC regulated growth. (Updated on 16-OCT-2003 to standardise OS field)
XX
XX SQ Sequence 2372 BP; 698 A; 491 C; 541 G; 642 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,97e-115 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 2 Gaps: 1

US-10-057-510-4 (1-216) x AAQ94589 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACTACTGATGCTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuSer 40
DB 372 GCTTCGGAACAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
DB 432 GTTGTGTGACAAAAGACATTAATCTACTATGAAGAGGTCTCTTTTATCTTGGCCAGTAT 491
QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
DB 492 ATTATGACTAAACGATTATATGATGAGAAAGCAACATATATGATATTTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
DB 552 CTTCTAGAGAGATTGTTTGGGTGCGCAAGCTTCTGTGAAAGAGACAGAGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
DB 612 ACCATGATCTACAGAACTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
DB 672 TCTGTGAGTGAGAACAGGTGTACCTTGAAGGTGGAGTGATCAAAAGGACCTTGTACAA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
DB 732 GAGCTTTCAGGAGAGAACCTTCATCTTCATTTGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
DB 792 AGAGGAGAGCAATATAGTGAGACAGAGAAATTCAGATGAATATCTGTTGTAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
DB 852 AGAAACCCCAATCTGATAGTATTTCCCTTTTCCCTTGGATGAAGCCCTGGCTCTGTGT 911
QY 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
DB 912 GTAATAAGGGAGATATGTTGTAAGAGAGCAGTAGCAGTGAATCTACAGG 962

RESULT 10
AAT45151

```

```

ID AAT45151 standard; cDNA; 2372 BP.
XX
XX AC AAT45151;
XX
XX DT 25-MAR-2003 (revised)
XX 28-JAN-1997 (first entry)
XX
XX DE Human MDM-2 gene cDNA clone, involved in tumour-development.
XX
XX KW p53; MDM-2; binding-inhibitor; identification; tumour; cancer; neoplasia;
XX antibody fusion protein; therapy; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX CDS 312..1787
XX PT /*tag= a
XX
XX XX US5550023-A.
XX
XX XX 27-AUG-1996.
XX
XX XX 18-MAY-1994; 94US-00245500.
XX
XX XX 07-APR-1992; 92US-00867840.
XX 23-JUN-1992; 92US-00903103.
XX 07-APR-1993; 93US-00044619.
XX
XX PA (UJJO ) UNIV JOHNS HOPKINS.
XX
XX PI Vogelstein B, Kinzler KW;
XX
XX XX WPI; 1996-401591/40.
XX DR P-PSDB; AAW07887.
XX
XX PT Identification of cpds. interfering with human MDM2/p53 binding - useful
XX as therapeutic agents to treat human neoplastic cells.
XX
XX PS Example 1; Col 21-26; 36pp; English.
XX
XX CC AAT45151 is a cDNA clone of the human MDM-2 gene derived from a human
XX colon carcinoma cell line CaCo-2. The MDM-2 protein produced by this
XX clone is used in a method for identifying compounds that interfere with
XX the binding of p53 and MDM-2. In binding the p53 protein, the MDM-2
XX protein releases a cell from p53-regulated growth, allowing cancers to
XX develop. Therefore compounds identified as interfering with the binding
XX of MDM-2 to p53 are potentially useful in the treatment of human
XX neoplastic cells. In the method pref. one or both of the proteins is a
XX fusion protein esp. with an antibody or antibody fragment which aids
XX separation and identification. (Updated on 25-MAR-2003 to correct PF
XX field.)
XX
XX SQ Sequence 2372 BP; 698 A; 491 C; 541 G; 642 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,97e-115 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 2 Gaps: 1

US-10-057-510-4 (1-216) x AAT45151 (1-2372)
QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
DB 312 ATGTGCAATACCAACATGCTGTACTACTGATGCTGTAAACCACTCACAGATTCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuSer 40
DB 372 GCTTCGGAACAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59

```

Db 432 GTTGGTGCACAAAGACACATTATCTACAAAGAGGTTCTTTTATCTTGGCCAGTAT 491
 QY 60 IleMetThrLysArgLeuValPheGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
 Db 492 ATTATGACTAAAGATTATATGATGAGAGCACACATATTTGATATTTGTTCAATGAT 551
 QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
 Db 552 CTTTACGAGATTGTTTGGCGTCCCAAGCTTCTCTGTGAAAGAGCACAGGAAATATAT 611
 QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnLysSerSerAspSerGlyThr 119
 Db 612 ACCATGATCTACAGGAACCTTGGTAGTCAATCAGCAGGAATATCGGACTCAGGTACA 671
 QY 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139
 Db 672 TCTGTGAGTGAGAACAGGTGTCACTTGAAGGTGGAGTGATCAAAAGGACCTTGTACAA 731
 QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
 Db 732 GAGCTTCAGGAAAGAGAAACCTTCATCTTCACATTTTGGTTTCTAGACCATCTACCTCATCT 791
 QY 160 ArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
 Db 792 AGNAGGAGGCAATTTAGTGAGACAGAGAAATTCAGATGAATTTCTGGTGAACACAA 851
 QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
 Db 852 AGAAACGCCACAAATCTCATAGTATTTCCCTTTCCTTTGATGAAAGCCCTGCTCTGTGT 911
 QY 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
 Db 912 GTAATAAGGGAGATATGTTGTGAAGAGACGTAGCAGTGAATCTACAGGG 962

RESULT 11

AAT66410

ID AAT66410 standard; cDNA; 2372 BP.

XX AAT66410;

XX 25-MAR-2003 (revised)

DT 14-SEP-2000 (revised)

DT 18-JUN-1997 (first entry)

XX Human MDM2 coding sequence.

XX Human; MDM2 protein; antibody; detection; cancer; diagnosis;

KW p53-regulated growth; ds.

XX Homo sapiens.

XX Key Location/Qualifiers

FH CDS 312..1787

FT /*tag= a

FT /product= "MDM2"

XX US5618921-A.

XX 08-APR-1997.

XX 17-FEB-1995; 95US-00390479.

XX 07-APR-1992; 92US-00867840.

XX 23-JUN-1992; 92US-00903103.

XX 07-APR-1993; 93US-00044619.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Vogelstein B, Kinzler KW, Burrell M, Hill DE;

XX WPI; 1997-225474/20.

XX P-PSDB; AAW15463.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Antibodies specific for human MDM2 protein - for diagnosis of cancer.
 XX Claim 1; Col 19-24; 35pp; English.
 XX This sequence encodes the human MDM2 protein. Antibodies that specifically bind to human MDM2 protein may be used for detecting elevated expression of the MDM2 gene in a human tissue or body fluid sample, esp. for cancer diagnosis. The antibodies may be used to interfere with the binding of p53 to MDM2. Elevated levels of MDM2 appear to sequester p53 and allow the cell to escape from p53-regulated growth. (N.B. Revised record issued to correct the sequence analysis field.)
 XX (Updated on 25-MAR-2003 to correct PF field.)

SQ Sequence 2372 BP; 698 A; 491 C; 541 G; 642 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,97e-115 Length: 2372
 Score: 1084.50 Matches: 216
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 0
 Query Match: 99.04% Indels: 1
 Gaps: 2

US-10-057-510-4 (1-216) x AAT66410 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrSerGlnIlePro 20
 Db 312 ATGTGCAATCAACATGTCTGTACCTACTGTGATGGTCTGTAAACCACTCAGATTTCCA 371

QY 21 AlaSerGluGlnThrLeuValAlaArgProLysProLeuLeuLysLeuLysLeu 40
 Db 372 GCITCGGAACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTATTAAAGTCT 431

QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu--GlnTyr 59
 Db 432 GTTGTGTCACAAAAGACACTTATCTATCAAGAGGTTCTTTTATCTTGGCCAGTAT 491

QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
 Db 492 ATTATGACTAAACGATTATATGATGAGAGCAACACATATTGTATATTCTCAATGAT 551

QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
 Db 552 CTTCTAGAGATTGTTTGGCGTCCCAAGCTTCTCTGTGAAGAGCACAGGAAATATAT 611

QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
 Db 612 ACCATGATCTACAGGAACCTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671

QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
 Db 672 TCTGTGAGTGAGAACAGGTGTCACTTGAAGGTGGGAGTGTCAAAAGGACCTTGTACAA 731

QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
 Db 732 GAGCTTCAGGAGAGAAACCTTCTATCTTCACATTTGGTTTCTAGACCATCTACCTCATCT 791

QY 160 ArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
 Db 792 AGAAGAGAGACAAATTAGTGAGACAGAGAAATTCAGATGAATTTCTGTGAAAGACCA 851

QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
 Db 852 AGAAACGCCACAAATCTGATAGTATTTCCTTTCTCTTTTGTAAAGCCCTGGCTCTGTGT 911

QY 200 ValIleArgGluIleCysCysGluArgSerSerSerGluSerThrGly 216
 Db 912 GTAATAAGGGAGATATGTTGTGAAGAGACGTAGCAGTGAATCTACAGGG 962

RESULT 12

AAT62065

ID AAT62065 standard; cDNA; 2372 BP.

```

XX AC AAT62065;
XX XX
XX DT 25-MAR-2003 (revised)
XX DT 05-JUN-1997 (first entry)
XX XX
XX XX Human MDM2 cDNA.
XX KW Human; MDM2; CaCo-2; colonic; carcinoma; probe; detection; amplification;
XX KW elevation; expression; diagnosis; neoplasia; neoplastic transformation;
XX KW sarcoma; colorectal; lung cancer; chronic myelogenous leukaemia; ds.
XX OS Homo sapiens.
XX XX
XX XX Key Location/Qualifiers
XX FH 312..1787
XX FT /*tag= a
XX FT /product= "MDM2"
XX XX
XX XX US5606044-A.
XX XX
XX XX 25-FEB-1997.
XX XX
XX XX 17-FEB-1995; 95US-00390546.
XX XX
XX XX 07-APR-1992; 92US-00867840.
XX PR 23-JUN-1992; 92US-00903103.
XX PR 07-APR-1993; 93US-00044619.
XX XX
XX XX (UYJO ) UNIV JOHNS HOPKINS.
XX XX
XX XX Kinzler KW, Vogelstein B, Hill DE, Burrell M;
XX XX
XX XX WPI; 1997-153623/14.
XX DR P-PSDB; AAW13380.
XX XX
XX XX Detection of amplification of human MDM2 gene - useful for diagnosis of
XX PT neoplasia or potential neoplastic transformation.
XX XX
XX XX Claim 1; Col 21-24; 35pp; English.
XX XX
XX XX The present sequence is the human MDM2 cDNA, which was isolated from a
XX CC human CaCo-2 colonic carcinoma cell cDNA library using a murine MDM2 cDNA
XX CC probe. The MDM2 cDNA can be used as a probe to detect the amplification
XX CC or elevated expression of a human MDM2 gene, which is diagnostic of
XX CC neoplasia or the potential for neoplastic transformation, useful for the
XX CC detection of, e.g. sarcomas, colorectal carcinoma, lung cancer and
XX CC chronic myelogenous leukaemia. (Updated on 25-MAR-2003 to correct PF
XX CC field.)
XX XX
XX XX Sequence 2372 BP; 698 A; 491 C; 541 G; 642 T; 0 U; 0 Other;
XX SQ
XX
Alignment Scores:
Pred. No.: 2.97e-115 Length: 2372
Score: 1084.50 Matches: 216
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 0
Query Match: 99.04% Indels: 1
DB: 2 Gaps: 1
XX
US-10-057-510-4 (1-216) x AAT62065 (1-2372)
XX
QY 1 MetCysAlaThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
Db 312 ATGTGCAATACCAACATGCTGTAACCTACTGATGCTGTACCACTCACAGATCCA 371
QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuSer 40
Db 372 GCTTCGGAACAAGAGACCCCTGGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431
QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
Db 432 GTTCGTGCACAAAAGACACTTATATGTAAGAGAGGTCTCTTTTATCTTGGCCAGTAT 491

```

```

QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
Db 492 ATTATGACTAAACGATTATATGATGAGAAGCAACAACATATTGTATATTGTTCAATGAT 551
QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
Db 552 CTTCTAGAGAGATTGTTTGGCGTCCCAAGCTTCTGTGAAAGAGCACAGGAAATATAT 611
QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGluSerSerAspSerGlyThr 119
Db 612 ACCATGATCTACAGGAACCTTGGTAGTAGTCAATCAGCAGGAATCATCGGACTCAGGTACA 671
QY 120 SerValSerGluAsnArgCysHisLeuGluGlySerAspGlnLysAspLeuValGln 139
Db 672 TCTGTGAGTGAGAACAGGTGTCACTTGAAGGTGGAGTGATCAAAAGGACCTTGTACAA 731
QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
Db 732 GAGCTTCAGGAAGAGAAACCTTCATCTTCATCTTGGTTTCTAGACCATCTACCTCATCT 791
QY 160 ArgArgArgAlaIleSerGluThrGluGlnSerAspGluLeuSerGlyGluArgGln 179
Db 792 AGAAGGAGAGCAATTTAGTGAGACAGAGAAATTCAGATGAATATCTGTGTAACGACAA 851
QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
Db 852 AGAAAACGCCCAAAATCTGATAGTATTTCCTTTTCCCTTTTGAAGAAAGCCTGGCTCTGTGT 911
QY 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
Db 912 GTATATAGGGAGATATGTTGTGAAGAGACAGTAGTAGTGAATCTACAGGG 962

```

```

RESULT 13
AAV20549
ID AAV20549 standard; cDNA; 2372 BP.
XX AC AAV20549;
XX DT 18-JUN-1998 (first entry)
XX DE Human MDM2 encoding cDNA.
XX KW Human; MDM2; hMDM2; tumour; cancer; diagnosis; neoplastic disease;
XX KW sarcoma; liposarcoma; malignant fibrous histiocytoma; osteosarcoma; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT 312..1787
XX FT /*tag= a
XX FT /product= "MDM2"
XX XX
XX XX US5736338-A.
XX XX
XX XX 07-APR-1998.
XX XX
XX XX 17-FEB-1995; 95US-00390517.
XX XX
XX XX 07-APR-1992; 92US-00867840.
XX PR 23-JUN-1992; 92US-00903103.
XX PR 07-APR-1993; 93US-00044619.
XX XX
XX XX (UYJO ) UNIV JOHNS HOPKINS.
XX XX
XX XX Vogelstein B, Kinzler KW, Hill DE, Burrell M;
XX DR WPI; 1998-239206/21.
XX DR P-PSDB; AAW48241.
XX XX
XX XX Cancer diagnosis - by determination of MDM2 protein.
XX XX
XX PS Claim 1; Col 21-24; 35pp; English.

```

XX The present sequence encodes human MDM2 (hMDM2) which is used in the
 CC method of the present invention. The present invention describes a method
 CC for diagnosing a neoplastic disease caused by overexpression of MDM2
 CC protein. The method comprises detecting an elevated cellular amount of
 CC this protein. The method is useful for the diagnosis of sarcoma,
 CC especially liposarcoma, malignant fibrous histiocyteoma or osteosarcoma
 XX
 SQ Sequence 2372 BP; 698 A; 491 C; 541 G; 642 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,97e-115 Length: 2372
 Score: 1084.50 Matches: 216
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 0
 Query Match: 99.04% Indels: 1
 DB: 2 Gaps: 1

US-10-057-510-4 (1-216) x AAV20549 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
 Db 312 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTAAACACCTCACAGATTCCA 371
 QY 21 AlaSerGluGlnGluThrValArgProLysProLeuLeuLeuLeuLeuLysSer 40
 Db 372 GCTTCGGAACAAGAGACCCCTGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431
 QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
 Db 432 GTTGTGGCAAAAGACACTTATCTACTATGAAGAGGTTCTTTTATCTTGGCCAGTAT 491
 QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
 Db 492 ATTATGACTAAACGATTATATGATGAGAAGCAACCAATATGTTATTTTCAATCAT 551
 QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
 Db 552 CTTCTAGGAGATTGTTTGGCGTGCACAGCTTCTCTGTGAAGAGACACAGGAAATATAT 611
 QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnGlnSerSerAspSerGlyThr 119
 Db 612 ACCATGATCTACAGGAATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 671
 QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
 Db 672 TCTGTGAGTGAGAACAGGTGTCACTTGAAGGTGGAGTGTCAAAAGGACCTTGTACAA 731
 QY 140 GluLeuGlnGluLysProSerSerSerHisLeuValSerArgProSerThrSerSer 159
 Db 732 GAGCTTCAGGAAGAGAAACCTTCATCTTCACATTTGGTTTCAGACCATCTACCTCATCT 791
 QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
 Db 792 AGAAGGAGAGCAATTAAGTGAGACAGAGAAATTCAGATGAATATCTGTTGAGACGACAA 851
 QY 180 ArgLysArgHisLysSerAspSerLysSerLysSerPheAspGluSerLeuAlaLeuCys 199
 Db 852 AGAAAACGCCCAAAATCTGATAGTATTTCCCTTTCCTTTGATGAAGCCTGGCTCTGTGT 911
 QY 200 ValIleArgGluIleCysCysGluArgSerSerSerSerGluSerThrGly 216
 Db 912 GTAATAAGGAGATATGTTGTAAGAAGACAGTAGACAGTGAATCTACAGGG 962

RESULT 14
 AAV28876
 ID AAV28876 standard; cDNA; 2372 BP.
 XX
 AC AAV28876;
 XX
 DT 10-AUG-1998 (first entry)
 XX Human MDM2 gene.

XX Human: p53; MDM2; tumour; growth inhibition; amplification;
 KW malignant fibrous histiocyteoma; liposarcoma; ds.
 XX

OS Homo sapiens.

Key Location/Qualifiers
 CDS 312..1787
 FT /*tag= a
 FT /product= "MDM2 protein"

XX US5756455-A.

XX 26-MAY-1998.

XX 17-FEB-1995; 95US-00390515.

XX 07-APR-1992; 92US-00867840.

XX 23-JUN-1992; 92US-00903103.

XX 07-APR-1993; 93US-00044619.

XX (UWJO) UNIV JOHNS HOPKINS.

XX Vogelstein B, Kinzler KW;

XX WPI; 1998-321574/28.

XX P-PSDB; AAW57241.

XX Inhibiting growth of tumour cells having MDM2 gene amplification - with

XX p53 protein fragment.

XX Example 1; Col 19-24; 40pp; English.

XX A method has been developed for inhibiting the growth of tumour cells
 CC containing a human MDM2 gene amplification. The method comprises treating
 CC the tumour cells with a DNA molecule that expresses a polypeptide capable
 CC of binding to human MDM2 protein. The present sequence represents the
 CC human MDM2 gene. The present invention describes three preferred
 CC polypeptides for binding human MDM2: (1) the polypeptide comprises amino
 CC acids 1-50 of p53 (see AAW57240); (2) the polypeptide comprises amino
 CC acids 13-41 of p53 (see AAW57240); and at least one additional p53
 CC residues on the N- or C-terminal side, provided that the polypeptide
 CC lacks the homooligomerisation domain of p53; (3) the polypeptide
 CC comprises amino acids 13-41 of p53 (see AAW57241) and at least nine
 CC additional p53 residues on the N- or C-terminal side, provided that the
 CC polypeptide lacks amino acids 138-393 of p53. Some malignant fibrous
 CC histiocyteomas and liposarcomas have an MDM2 gene amplification, so
 CC detection of increased expression of MDM2 gene products indicates
 CC tumorigenesis

XX SQ Sequence 2372 BP; 698 A; 491 C; 541 G; 642 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,97e-115 Length: 2372
 Score: 1084.50 Matches: 216
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 0
 Query Match: 99.04% Indels: 1
 DB: 2 Gaps: 1

US-10-057-510-4 (1-216) x AAV28876 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20

Db 312 ATGTGCAATACCAACATGCTGTACCTACTGATGCTGTAAACACCTCACAGATTCCA 371

QY 21 AlaSerGluGlnGluThrValArgProLysProLeuLeuLeuLysLysSer 40

Db 372 GCTTCGGAACAAGAGACCCCTGTTAGACCAAGCCATTGCTTTTGAAGTTATTAAAGTCT 431

QY 41 ValGlyAlaGlnLysAspThrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59

Db 432 GTTGTGGTGCACAAAGACACTTATCTATGAAAGAGGTTCTTTTATCTTGGCCAGTAT 491

QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnSer 79
 Db 492 ATTATGACTAAACGATTATATGATGAGAACCAACATATTTGTTCAATGAT 551
 QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
 Db 552 CTTCTAGGAGATTGTTGGCGCCAAAGCTTCTCTGTAAGACGACAGGAAATATAT 611
 QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnLysSerAspSerGlyThr 119
 Db 612 ACCATGATCTACAGGAACCTTGGTAGTAGTCAATCAGCAGGAATCATCGACTCAGGTACA 671
 QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
 Db 672 TCTGTGAGTGAGAACAGGTGTACCTTGAAGGTGGAGTGATCAAAAGGACCTTGTACAA 731
 QY 140 GluLeuGlnGluLysProSerSerHisLeuValSerArgProSerThrSerSer 159
 Db 732 GAGCTTCAGAGAGAACCTTCACTCTCACAATTTGGTTTCTAGACCATCTACCTCATCT 791
 QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
 Db 792 AGAAGGAGAGCAATTAGTCAGACAGAGAAATTCAGATGAATATCTGTGTAACGACAA 851
 QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
 Db 852 AGAAAACGCCCAAAATCTGATGATTTCCCTTTCCCTTTGATGAAGACCTGCTGTGT 911
 QY 200 ValIleArgGluLeuCysCysGluArgSerSerSerSerGlyThrGly 216
 Db 912 GTAATAAGGAGATATGTTGTGAAGAGACAGTACAGTGAATCTACAGGG 962

RESULT 15

AAV04836
 ID AAV04836 standard; cDNA; 2372 BP.

XX AAV04836;

DT 30-APR-1998 (first entry)

XX cDNA sequence of human MDM2.

XX MDM2; tumour; diagnosis; neoplasia; DNA binding protein; p53 polypeptide;
 KW binding; tumour cell; p53-regulated growth; inhibition;
 KW anti-cancer agent; ds.

XX Homo sapiens.

OS Key Location/Qualifiers
 FH 312..1787
 FT CDS /*tag= a

XX US5708136-A.

XX 13-JAN-1998.

XX 17-FEB-1995; 95US-003990516.

XX 07-APR-1992; 92US-00867840.

XX 23-JUN-1992; 92US-00903103.

XX 07-APR-1993; 93US-00044619.

XX (UWJO) UNIV JOHNS HOPKINS.

XX Vogelstein B, Kinzler KW, Burrell M, Hill DE;

XX WPI; 1998-100408/09.

XX Human MDM2 binding polypeptide - comprises fragments of p53, useful in re-
 XX -establishing p53-regulated growth control in cells over-expressing MDM2.

XX Disclosure; Col 19-24; 41pp; English.

XX

CC The present sequence encodes human MDM2. The MDM2 gene is amplified in
 CC some human tumours. The amplification of this gene is diagnostic of
 CC neoplasia or its potential. It is speculated that the MDM2 protein is a
 CC potential DNA binding protein that functions in the modulation of
 CC expression of other genes and, when present in excess, interferes with
 CC normal constraints on cell growth. A cell containing three recombinant
 CC DNA constructs was produced. These constructs encode an MDM2 protein
 CC fused to a sequence-specific DNA binding domain, a p53 polypeptide fused
 CC to a transcriptional activation domain, and a reporter gene downstream
 CC from a DNA element which is recognised by the sequence-specific DNA-
 CC binding domain. The cell is used to identify a compound which interferes
 CC with the binding of MDM2 and p53. Since MDM2 is overexpressed in tumour
 CC cells and since binding of MDM2 to p53 appears to allow tumour cells to
 CC escape from p53-regulated growth, compounds that inhibit such binding
 CC would be useful as anti-cancer agents

SQ Sequence 2372 BP; 698 A; 491 C; 541 G; 642 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.97e-115 Length: 2372
 Score: 1084.50 Matches: 216
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 0
 Query Match: 99.04% Indels: 1
 DB: 2 Gaps: 1

US-10-057-510-4 (1-216) x AAV04836 (1-2372)

QY 1 MetCysAsnThrAsnMetSerValProThrAspGlyAlaValThrThrSerGlnIlePro 20
 Db 312 ATGTGCAATACCAACATGTCTGTACCTACTGATGGTCTGTACACCTCACAGTTCCA 371
 QY 21 AlaSerGluGlnGluThrLeuValArgProLysProLeuLeuLeuLeuLeuLeuSer 40
 Db 372 GCTTCGGACACAGAGACCCCTGTTAGACCAAGCCATTGCTTTTGAAGTATTAAAGTCT 431
 QY 41 ValGlyAlaGlnLysAspThrTyrThrMetLysGluValLeuPheTyrLeu---GlnTyr 59
 Db 432 GTTGTGTGCACAAAAGACACTTATCTATGAAGAGGTCTTTTATATCTTGGCCAGTAT 491
 QY 60 IleMetThrLysArgLeuTyrAspGluLysGlnGlnHisIleValTyrCysSerAsnAsp 79
 Db 492 ATTATGACTAAACGATTATATGATGAGAGCAACACATATTTGTTATTTGTTCAATGAT 551
 QY 80 LeuLeuGlyAspLeuPheGlyValProSerPheSerValLysGluHisArgLysIleTyr 99
 Db 552 CTTCTAGGAGATTGTTTGGCGTGCACAGCTTCTCTGTGAAGAGACACAGAAAATATAT 611
 QY 100 ThrMetIleTyrArgAsnLeuValValAsnGlnGlnLysSerAspSerGlyThr 119
 Db 612 ACCATGATCTACAGGAACCTTGTAGTAGTCAATCAGCAGGAATCATCGACTCAGGTACA 671
 QY 120 SerValSerGluAsnArgCysHisLeuGluGlyGlySerAspGlnLysAspLeuValGln 139
 Db 672 TCTGTGAGTGAGAACAGGTGTACCTTGAAGGTGGAGTGATCAAAAGGACCTTGTACAA 731
 QY 140 GluLeuGlnGluLysProSerSerHisIleValSerArgProSerThrSerSer 159
 Db 732 GAGCTTCAGGAAGAGAAACCTTCACTCTCACAATTTGGTTTCTAGACCATCTACCTCATCT 791
 QY 160 ArgArgArgAlaIleSerGluThrGluGluAsnSerAspGluLeuSerGlyGluArgGln 179
 Db 792 AGAAGGAGAGCAATTAGTGAGACAGAGAAATTCAGATGAATATCTGTGTGAACGACAA 851
 QY 180 ArgLysArgHisLysSerAspSerIleSerLeuSerPheAspGluSerLeuAlaLeuCys 199
 Db 852 AGAAAACGCCCAAAATCTGATGATTTCCCTTTCCCTTTGATGAAGACCTGCTGTGT 911
 QY 200 ValIleArgGluLeuCysCysGluArgSerSerSerSerGlyThrGly 216
 Db 912 GTAATAAGGAGATATGTTGTGAAGAGACAGTACAGTGAATCTACAGGG 962

Search completed: March 14, 2004, 22:06:03
Job time : 365.517 secs